PHD

Entropic Gradient Flows on the Wasserstein Space via Large Deviations from Thermodynamic Limits

Laschos, Vaios

Award date:
2013

Awarding institution:
University of Bath

Entropic Gradient flows on the Wasserstein space via large deviations from thermodynamic limits

submitted by

Vaios Laschos

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Mathematical Sciences

April 2013

Summary

In a seminal work, Jordan, Kinderlehrer and Otto proved that the Fokker-Planck equation can be described as a gradient flow of the free energy functional in the Wasserstein space, bringing this way the statistical mechanics point of view on the diffusion phenomenon to the foreground. The aim of this thesis is to show that it is possible to retrieve this natural coupling of functional and metric, by studying the large deviations of particle models. More specifically, for the case where the ambient space is the real line, it is proved that the free energy functional can be retrieved as an asymptotic Gamma-limit ($\tau \to 0$) of the rate function of a large deviation principle, minus the square of the Wasserstein distance (normalized by time). Furthermore, for a special case where both measures in the definition of the rate function are Gaussians, its value and the rate of convergence are being calculated explicitly.

CONTENTS

LIST OF NOTATIONS

CHAPTER 1

INTRODUCTION

Diffusion phenomena are a class of transport phenomena that occur in a wide range of sciences. The linear partial differential equation (PDE)

$$\partial_t \mu(t) = \Delta\mu(t) + div(\mu(t)\nabla\Psi), \quad \mu(0) = \mu_0 \in \mathcal{P}_2(\mathbb{R}^d) \tag{1.0.1}$$

is called the Fokker-Planck equation and it characterizes a big class of diffusion phenomena. When the potential Ψ is zero the equation takes the following simple form

$$\frac{\partial\mu(t)}{\partial t} = \Delta\mu(t), \qquad \mu(0) = \mu_0 \tag{1.0.2}$$

which is the basic model for diffusion and dates back even before 1822, when Fourier in his work titled *Thorie analytique de la chaleur* (The Analytic Theory of heat) studied and solved it. The derivation of the equation was based on the simple assumption that locally, the rate at which heat spreads (heat flux) is proportional to the difference of the temperature for two adjusted areas . Couple of years later Flick suggested a similar principle for the diffusion of fluids [1].

In 1905, Albert Einstein in an attempt to explain the phenomenon of Brownian motion observed by Robert Brown in 1827, derived the diffusion equation following a totally different line of reasoning. Motivated by the, still in development at the time, atomistic theory, he assumed that the motion is due to a constant

[1] Other similar principles are Darcy's law (hydraulic flow) and Ohm's law (charge transport)

bombardment by smaller particles (atoms). With that in mind he posed some "restrictions" that the motion should obey, like being homogeneous in time and space. Also he asked that the motion will satisfy the experimental findings, that in average the square of displacement is proportional to the time. He reached this way the conclusion that this is possible only if the densities of the particles satisfy the diffusion equation.

Today there are many derivations of the (1.0.2) apart from Fourier's Law. The same holds for the more general (1.0.1). For example it is well know that (1.0.1) can be described as the gradient flow of the Dirichlet functional in $\mathcal{L}^2(\mathbb{R}^d)$. In fact, one can find an infinite number of couples of spaces X and functionals G so that (1.0.1) can be described as a gradient flow of G in X. However, some of them are more natural to interpret than others.

In physics, it is well known that the free energy functional $\mathcal{F} = \mathcal{S} + \mathcal{E}$, where

$$\mathcal{S}(\mu) = \begin{cases} \displaystyle\int \frac{d\mu}{d\mathbf{r}} \log(\frac{d\mu}{d\mathbf{r}}) d\mathbf{r} & \text{when } \mu \ll \mathcal{L} \\ \infty & \text{otherwise,} \end{cases} \tag{1.0.3}$$

is the entropy functional and

$$\mathcal{E}(\mu) = \int_{\mathbb{R}^d} \Psi(\mathbf{r})\mu(d\mathbf{r}),$$

is the potential energy, is decreasing along solutions of (1.0.1). Therefore, it can be argued that the free energy acts as a "driving force" for the phenomenon, which seems very intuitive when someone thinks of the particle interpretation of diffusion. A natural question arises. Does it exist a metric space X where the Fokker-Planck (1.0.1) equation is the gradient flow of the free energy $\mathcal{F}$ with respect to that space X? In [19] and [21], the authors show that this holds for the Wasserstein space $(\mathcal{P}_2(\mathbb{R}^d), W_2)$.

By using the following iteration scheme

$$\mu(t_k) = \arg\min \left\{ \frac{W_2^2(\mu(t_{k-1}), \mu)}{2\tau} + \mathcal{F}(\mu) \right\}, \quad t_k = t_{k-1} + \tau, \quad \text{and} \quad \mu(0) = \mu_0$$

for some initial data $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, and by interpolating constantly, they formed a family of curves indexed by τ. In the sequel, they proved that the family converges to a limit curve when τ goes to zero and that the limit curve is a solution to the Fokker-Planck equation (1.0.1) with initial data μ_0.

This very important result originated by an observation by Felix Otto, namely that the entropic gradient with respect to the pseudo-Riémannian structure (see chapter 5) in Wasserstein space is the Laplacian. This gives rise to following question. Was there a way to derive such a "natural" couple of space and functional, or even better, *is it possible to find a systematic method that gives such couples of metric spaces and functionals, in order to describe the solution of a PDE which arises as thermodynamic limit of a particle model?*

The authors of [20] and sequentially of [1] suggest that such a method could arise from the theory of large deviations. It is known that the empirical measure $L_N(\tau)$ that corresponds to the model of N independent Brownian particles with initial positions $(r_{i,0})_{1\le i\le N}$ with $\frac{1}{N}\sum_{i=1}^{N}\delta_{r_{i,0}} \to \mu_0$, not only converges weakly to the solution $\mu(\tau)$ of (1.0.2), but also satisfies a Large Deviation Principle (LDP) with rate function

$$J_\tau(\mu|\mu_0) = \inf\left\{\mathcal{H}(\xi\|\mu_0 P_\tau) : \xi \in \Pi(\mu_0, \mu)\right\}, \tag{1.0.4}$$

where $\mathcal{H}$ is the relative entropy

$$\mathcal{H}(\xi'\|\xi) = \begin{cases} \displaystyle\int \frac{d\xi'}{d\xi}\log(\frac{d\xi'}{d\xi})d\xi & \text{when } \ \xi' << \xi \\ \infty & \text{otherwise,} \end{cases} \tag{1.0.5}$$

and P_τ is the fundamental solution of the diffusion equation. In [20] the author proves that $\tau J_\tau(\cdot|\mu_0)$ Gamma-converges to the Wasserstein metric $W_2(\cdot,\mu_0)$ as τ goes to 0, recovering this way the Wasserstein distance from $J_\tau(\cdot|\mu_0)$. However, as it became apparent in [1], it is not only possible to obtain the appropriate space by using $J_\tau(\cdot|\mu_0)$ but the right functional $\mathcal{F}$ can be retrieved as well.

In [1], the authors conjectured that it is possible to retrieve the free energy functional by a Gamma-limit. They claimed that the following holds.

$$J_\tau(\cdot\,|\mu_0) - \frac{W_2^2(\mu_0,\,\cdot\,)}{4\tau} \to \frac{1}{2}\mathcal{F}(\,\cdot\,) - \frac{1}{2}\mathcal{F}(\mu_0), \qquad \text{in } \mathcal{P}_2(\mathbb{R}^d), \tag{1.0.6}$$

as $\tau \to 0$, where the convergence is Gamma-convergence. In the same paper, a very special modification of the above statement was proved. Instead of the real line, the flat torus was taken as the ambient space. Even more not all measures were being considered, but only those that are sufficiently close to the Lebesgue measure. Finally the potential Ψ it was assumed to be equal to zero.

1.1 Outline

The aim of this thesis is to extend the results from [1]. The first result is for the case where $\Psi = 0$ and both the set of initial data and the domain of J_τ are identified with the set of the one dimensional Gaussians. More specifically

Theorem 1.1.1. *Let $\mathcal{N}(m_0, \sigma_0^2)$ be a normal distribution. Then for the rate functional J_τ it holds that*

$$J_\tau(\cdot; \mathcal{N}(m_0, \sigma_0^2)) - \frac{W_2^2(\cdot, \mathcal{N}(m_0, \sigma_0^2))}{4\tau} \to \frac{1}{2}\mathcal{S}(\cdot) - \frac{1}{2}\mathcal{S}(\mathcal{N}(m_0, \sigma_0^2)),$$

locally uniform and in the sense of Γ-convergence with respect to the weak topology, on the submanifold of the Gaussians.

In the second part, by using a different approach that involves large deviations for trajectories, we treat the more general case where Ψ can be nonzero and the set of the initial data as well the domain of J_τ are much larger.

Theorem 1.1.2. *Let $\mu_0 \in \mathcal{P}_2(\mathbb{R})$ be absolutely continuous with respect to the Lebesgue measure and with density $\rho_0(x)$ being bounded from below by a positive constant in every compact set. Assume that $\int_{\mathbb{R}} |\nabla\Psi(x)|^2 \, \mu_0(dx)$ and the Fisher information $I(\mu_0)$ are finite, and that Ψ satisfies Assumption 6.2.1. Then we have*

$$J_\tau(\cdot\,|\mu_0) - \frac{W_2^2(\mu_0, \cdot\,)}{4\tau} \xrightarrow[\tau\to 0]{\Gamma} \frac{1}{2}\mathcal{F}(\,\cdot\,) - \frac{1}{2}\mathcal{F}(\mu_0), \quad \text{in} \quad \mathcal{P}_2(\mathbb{R}). \tag{1.1.1}$$

As it was mentioned before, the type of convergence in the main result is Gamma-convergence. The definition of Gamma convergence is provided for the convenience of the reader.

Definition 1.1.3. *Let X be a topological space that satisfies the first axiom of countability. We say that F_n* Γ-converges *to F if the following conditions are satisfied:*

1. *(Lower bound) For every $x \in X$ and for every sequence $\{x_n\}_{n\in\mathbb{N}}$ converging to x in X,*
$$F(x) \leq \liminf_{n\to\infty} F_n(x_n).$$

2. *(Recovery sequence) For every $x \in X$, there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ converging to x in X such that*
$$F(x) = \lim_{n\to\infty} F_n(x_n).$$

We then write $F_n \xrightarrow{\Gamma} F$.

Unlike what the title of the thesis may indicates, it is not directly proven that the Fokker-Planck equation (1.0.1) is a gradient flow of the free energy on the Wasserstein space, by studying the large deviations of the particle models. What is mainly proven, is that the "building elements" can be retrieved. However by applying a more naive approach, someone can rewrite J_τ as $\frac{W_2^2(\mu_0,\cdot)}{4\tau} + \frac{1}{2}\mathcal{F}(\cdot) - \frac{1}{2}\mathcal{F}(\mu_0)$ for very small values of τ and claim that the minimizers of both functionals are close. Therefore by looking for minimizers of $\frac{W_2^2(\mu_0,\cdot)}{4\tau}+\frac{1}{2}\mathcal{F}(\cdot)-\frac{1}{2}\mathcal{F}(\mu_0)$ someone expects to get the solution to the Fokker-Planck equation after time step τ. On the other hand, the minimizers of $\frac{W_2^2(\mu_0,\cdot)}{4\tau} + \frac{1}{2}\mathcal{F}(\cdot) - \frac{1}{2}\mathcal{F}(\mu_0)$ are giving rise to gradient flows of the free energy to the Wasserstein space and this is how the indirect connection is made.

One of the main reasons that the first result is included although the second one is stronger, is that for the case of Gaussians the minimizers are calculated explicitly, as well the value of (1.0.4). Another reason is that different techniques are used. It is worth mentioning that the last couple of years, while this thesis was written, the authors of [1], along with close collaborators, proved similar results for other couples of spaces and functionals([15],[22]).

In the second chapter, the notion of the Wasserstein space is introduced. Historical background is given and some of the basic properties are explained. Furthermore, the concept of gradient flows in general metric spaces is described. After all the machinery has been introduced, the chapter concludes with a "summary" of [19].

The third chapter provides a very brief introduction to the theory of large deviations. The model of independent Brownian particles in a potential is discussed and the rate function for the model is introduced.

In the fourth chapter, the (1.0.6) for the case of Gaussians on the real line is explained in more detail and some of its implications are highlighted. Furthermore, its validity for the case where both the set of initial data and the domain of J_τ is the set of the one dimensional Gaussians is proven.

In chapter four the Wasserstein space is revisited. The differential structure heuristically introduced by Felix Otto and rigorously developed by Ambrosio, Savare and Gigli [3] is studied. Furthermore some functionals on the Wasserstein are introduced and investigated.

Chapter five is a continuation of chapter two, where large deviations on trajectories are studied instead of large deviations on time points. By connecting rate functions via the contraction principle we get a formula for J where the free energy appears explicitly.

Chapter six conludes the thesis with the proof of the most general case of (1.0.6), and some discussion regarding the result.

CHAPTER 2

WASSERSTEIN SPACE

2.1 The Monge-Kantorovich problem

The Wasserstein space has its origins on a problem posed by Gaspard Monge in 1781, regarding the transportation of goods. In mondern language, let μ be a measure in $\mathbb{R}^d$ that represents the supply of a product and ν another measure of equal total mass that characterizes the demand. Let also $C(\mathbf{r}, \mathbf{r}') \colon \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty)$ be a function that represents the cost of moving a single item from $\mathbf{r}$ to $\mathbf{r}'$. The question posed by Monge is the following

Question 2.1.1. *Does a map $\mathbf{T}^\star$ from $\mathbb{R}^d$ to $\mathbb{R}^d$, that describes which supplier sends to which "customer", covers all the demand and minimizes the cost of the transportation over all such maps exist?*

Mathematically, the assumption that $\mathbf{T}$ covers the demand, it can be formalized with the notion of the pushforward measure.

Definition 2.1.2. *Let $\mathbf{T} \colon \mathbb{R}^d \to \mathbb{R}^d$ a Borel map, and $\mu \in \mathcal{P}(\mathbb{R}^d)$ then $\mathbf{T}_\# \mu \in \mathcal{P}(\mathbb{R}^d)$ defined by*

$$\mathbf{T}_\# \mu(E) = \mu(\mathbf{T}^{-1}(E)), \quad \forall E \subset Y \tag{2.1.1}$$

is called the push forward of μ through $\mathbf{T}$.

With $S(\mu, \nu)$ the set of all maps $\mathbf{T}$ such that $\nu = T_{\#}\mu$ is denoted. Now the question can be reformulated as follows.

Question 2.1.3. *Is it possible to find a Borel map* $\mathbf{T}^\star : \mathbb{R}^d \to \mathbb{R}^d$ *such that* $\nu = \mathbf{T}^\star{}_{\#}\mu$ *and*

$$\int_{\mathbb{R}^d} c(\mathbf{r}, \mathbf{T}^\star(\mathbf{r}))d\mu = \inf_{\mathbf{T} \in S(\mu,\nu)} \int_{\mathbb{R}^d} c(\mathbf{r}, \mathbf{T}(\mathbf{r}))d\mu \ \ ?$$

It is easy to see that this question is ill-posed in many ways. To start, if one considers the example where there exists only one supplying location with two products and two locations each demanding one product, will immediately realize that there is no way to describe any reasonable solution with just the help of a map (i.e $S(\mu, \nu) = \varnothing$).

Furthermore, in general the set $S(\mu, \nu)$ has neither nice convexity properties nor is it compact with respect to some "natural" weak topology, making it impossible to use the direct method in the calculus of variation.

Example 2.1.4 ([26])**.** *Let* $\mu = \nu = \mathcal{L}_{|_{[0,1]}}$. *Let also* $T_1(x) = x$ *and* $T_2(x) = \min\{2x, 2 - 2x\}$. *Obviously* $\nu = T_{1\#}\mu$ *and* $\nu = T_{2\#}\mu$, *but at the same time* $\nu \neq (\frac{2}{3}T_1 + \frac{1}{3}T_2)_{\#}\mu$, *therefore* $S(\mathcal{L}_{|_{[0,1]}}, \mathcal{L}_{|_{[0,1]}})$ *is not convex.*

Actually, for a suitable choice of μ, ν, it is possible to construct a sequence T_n of optimal plans that has a unique weak limit (weak $L^p, 1 \leq p \leq \infty$) that it is not a plan (see [2, Page 4] or [26, Page 3]).

In 1942 Leonid Kantorovich proposed a relaxed version of Monge's problem. Contrary to Monge, Kantorovich allowed the mass at one point to split and be distributed to several locations. To say it in a more mathematical way, he searched for solutions not among maps that push a measure $\mu \in \mathcal{P}(\mathbb{R}^d)$ forward to a measure $\nu \in \mathcal{P}(\mathbb{R}^d)$ but among all transportation plans ξ (i.e., measures in $\mathbb{R}^d \times \mathbb{R}^d$) with $(\pi_0)_{\#}\xi = \mu$ and $(\pi_1)_{\#}(\xi) = \nu$, where $\pi_{i_1,i_2,\ldots,i_k} \colon \mathbb{R}^{nd} \to \mathbb{R}^{kd}$ with $\pi_{i_1,i_2,\ldots,i_k}(\mathbf{x}_0, \ldots, \mathbf{x}_{N-1}) = (\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_k})$. In the sequel we will denote with $\Pi(\mu, \nu)$ the set of all measures $\xi \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ with $(\pi_0)_{\#}(\xi) = \mu$ and $(\pi_1)_{\#}(\xi) = \nu$.

In a rigorous way the problem can be restated as follows.

Question 2.1.5. *Is there a* $\xi^\star \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d)$ *such that*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} c(\mathbf{r},\mathbf{r}')d\xi^\star = \inf_{\xi\in\Pi(\mu,\nu)} \int_{\mathbb{R}^d\times\mathbb{R}^d} c(\mathbf{r},\mathbf{r}')d\xi \ \ ?$$

In the sequel we will assume that c is of the form $c(\mathbf{r},\mathbf{r}') = c'(\mathbf{r}-\mathbf{r}')$ for some convex function c' and even more, our main object of study will be for $c(\mathbf{r},\mathbf{r}') = \|\mathbf{r}-\mathbf{r}'\|^2$.

It is easy to see that $\Pi(\mu,\nu)$ is convex and compact with respect to the narrow topology. We remind the reader of the notion of narrow convergence.

Definition 2.1.6. *A sequence* $\{\mu_n\}_{n\in\mathbb{N}} \subset \mathcal{P}(\mathbb{R}^d)$ *is narrowly convergent to* $\mu \in \mathcal{P}(\mathbb{R}^d)$ *as* $n \to +\infty$ *if*

$$\lim_{n\to+\infty} \int_{\mathbb{R}^d} f(\mathbf{r})d\mu_n = \int_{\mathbb{R}^d} f(\mathbf{r})d\mu$$

for all $f \in C_b^0(\mathbb{R}^d)$.

To see why $\Pi(\mu,\nu)$ is compact, we will need the notion of tightness along with a well-known theorem of Prokhorov.

Definition 2.1.7. *A set* $\mathcal{K} \subset \mathcal{P}(\mathbb{R}^d)$ *is* tight *if for every* $\epsilon > 0$, *there exists a compact set* K_ϵ *such that for all* $\mu \in \mathcal{K}$ *we have* $\mu(\mathbb{R}^n \setminus K_\epsilon) \leq \epsilon$.

Theorem 2.1.8 (Prokhorov,[6])**.** *If* $\mathcal{K}$ *is tight, then it is relatively compact in* $\mathcal{P}(\mathbb{R}^d)$ *endowed with the narrow convergence topology. Conversely, every relatively compact set in* $\mathcal{P}(\mathbb{R}^d)$ *endowed with the narrow convergence topology is tight.*

Now if for arbitrary ϵ, μ and ν we pick $K_{\mu,\epsilon}$ and $K_{\nu,\epsilon}$ respectively as in the above definition, and we form $K_{\mu,\epsilon}\times K_{\nu,\epsilon}$, it is easy to see that $\xi(\mathbb{R}^d\times\mathbb{R}^d - K_{\mu,\epsilon}\times K_{\nu,\epsilon}) < 2\epsilon$ for every $\xi \in \Pi(\mu,\nu)$ and therefore

Theorem 2.1.9 ([2])**.** *Let* $\mu,\nu \in \mathcal{P}(\mathbb{R}^d)$. *The set* $\Pi(\mu,\nu)$ *is compact with respect to the narrow topology.*

Also it is easy to check that the functional $\mathcal{I}_c(\xi) = \int_{\mathbb{R}^d\times\mathbb{R}^d} c(\mathbf{r},\mathbf{r}')d\xi$, which we will call the *total cost* for the transportation plan ξ, is lower semi continuous.

Now since $\Pi(\mu, \nu)$ is compact and the functionals $\int_{\mathbb{R}^d \times \mathbb{R}^d} c(\mathbf{r}, \mathbf{r}')d\xi$ is lower semi continuous and bounded from below, a minimizer exists.

To repeat a standard argument for the theory of calculus of variations, let ξ_n be a sequence such that $\mathcal{I}_c(\xi_n) \to \inf_{\xi \in \Pi(\mu,\nu)} \mathcal{I}_c(\xi)$. Now by compactness of $\Pi(\mu, \nu)$, we deduce that it exists subsequence ξ_{k_n} such that ξ_{k_n} converges to some ξ_0. Now by lower semi continuity we have

$$\inf_{\xi \in \Pi(\mu,\nu)} \mathcal{I}_c(\xi) \leq \mathcal{I}_c(\xi_0) \leq \lim_{n \to \infty} \mathcal{I}_c(\xi_{k_n}) = \inf_{\xi \in \Pi(\mu,\nu)} \mathcal{I}_c(\xi)$$

and therefore

$$\mathcal{I}_c(\xi_0) = \inf_{\xi \in \Pi(\mu,\nu)} \mathcal{I}_c(\xi).$$

The set of all optimal plans from μ to ν will be denoted by $\Pi_c^\star(\mu, \nu)$ or $\Pi^\star(\mu, \nu)$ when c is obvious.

2.2 The Wasserstein space

As it will became apparent in the sequel, one of the most interesting applications of the Monge-Kantorovich problem is for the case where the cost function C is equal to the square of the Euclidean distance. If we choose μ, ν in

$$\mathcal{P}_2(\mathbb{R}^n) = \left\{ \mu \in \mathcal{P}(\mathbb{R}^n) : \int |\mathbf{r}|^2 \, d\mu < \infty \right\},$$

it is easy to see that the transportation cost of moving μ to ν is finite. It is also trivial to check that it satisfies the coincidence axiom and the symmetric property of a metric. Actually when the cost is convex, the triangular inequality also holds and indeed the optimal transportation cost between two measures gives rise to a metric. In our case, where the cost function is the square of the distance, the metric is called *2-Wasserstein distance* were the term "was coined by R.L. Dobrushin in 1970, after the Russian mathematician Leonid Nasonovich Vasershtein who introduced the concept in 1969" [1]. It is defined by

$$W_2^2(\mu,\nu) = \min_{\xi\in\Pi(\mu,\nu)} \left\{ \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} |\mathbf{r}-\mathbf{r}'|^2 d\xi \right\}.$$

The following theorem gives a very useful equivalent definition for converging sequences in the Wasserstein space.

Theorem 2.2.1 ([28])**.** *Let (μ_n) be a sequence in $\mathcal{P}_2(\mathbb{R}^d)$. Let also $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, then*

$$\mu_n \to \mu \text{ as } n\to\infty \quad \text{if and only if} \quad \begin{array}{ll} (i) & \mu_n \to \mu \text{ narrowly, and} \\ (ii) & \displaystyle\int_{\mathbb{R}^d} |\mathbf{r}|^2\, d\mu_n \to \int_{\mathbb{R}^d} |\mathbf{r}|^2\, d\mu. \end{array}$$

Although we know that the Kantorovich problem always has a solution for two measures μ,ν in the Wasserstein space, there still remains the question of when the initial Monge problem has one solution too. Equivalently, when does a plan $\mathbf{T}^\star \in S(\mu,\nu)$ exist, such that a minimizer $\xi^\star \in \Pi^\star(\mu,\nu)$ is of the form $\xi^\star = (\mathbf{I}\times\mathbf{T}^\star)_{\#}\mu$, where I is the identity map?

Before we proceed, we are going to provide some definitions.

Definition 2.2.2 (c - c hypersurfaces)**.** *A set $E\subset\mathbb{R}^d$ is called a* c - c hypersurface *if, in a suitable system of coordinates, it is the graph of the difference of two real valued convex functions, i.e. if there exists convex functions $f,g:\mathbb{R}^{d-1}\to\mathbb{R}$ such that*

$$E = \left\{(t,y)\in\mathbb{R}^d : y\in\mathbb{R}^{d-1}, t\in\mathbb{R}, t=f(y)-g(y)\right\}$$

Definition 2.2.3 (Regular measures on $\mathbb{R}^d$)**.** *A measure $\mu\in\mathcal{P}_2(\mathbb{R}^d)$ is called* regular *if $\mu(E)=0$ for any c - c hypersurface $E\subset\mathbb{R}^d$.*

The set of all regular measures in $\mathbb{R}^d$ will be denoted with $\mathcal{P}_{2,r}(\mathbb{R}^d)$. Now we can state a result concerning existence and uniqueness of optimal maps:

Theorem 2.2.4 (Brenier)**.** *Let $\mu\in\mathcal{P}_2(\mathbb{R}^d)$. Then the following are equivalent:*

1. *For every $\nu\in\mathcal{P}_2(\mathbb{R}^d)$, there exists only one optimal transport plan from μ to ν and this plan is induced by a map $\mathbf{T}^\star$,*

[1]Wikipedia (`http://en.wikipedia.org/wiki/Transportation_theory_(mathematics)`)

2. μ is regular.

If either 1, or 2 hold, the optimal map $\mathbf{T}^\star$ can be recovered by taking the gradient of a convex function.

Observe that absolutely continuous measures are automatically regular. Therefore we get the following statement.

Corollary 2.2.5. *If $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ is absolutely continuous, then for every $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ there exists a $\mathbf{T}^\star \in S(\mu,\nu)$ such that the minimizer $\xi^\star \in \Pi^\star(\mu,\nu)$ in the definition of the Wasserstein distance is of the form $\xi^\star = (\mathbf{I} \times \mathbf{T}^\star)_{\#}\mu$.*

2.3 Geodesics

Definition 2.3.1 (Constant speed geodesics). *Let (X,d) a metric space. A curve $x\colon [0,1] \to X$ is a* geodesic *if*

$$d(x(t_2),x(t_1)) = d(x(0),x(1))(t_2 - t_1), \quad \forall 0 \le t_1 \le t_2 \le 1.$$

The Wasserstein space is a geodesic space, which means that for every choice of measures $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ it exists a constant speed geodesic connecting them. Moreover we have the following nice theorem that characterizes the geodesics.

Theorem 2.3.2. *Let $\mu_0, \mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\xi^\star \in \Pi^\star(\mu_0,\mu_1)$. The curve*

$$t \to \mu(t) := \mu_{0\to 1}(t) = ((1-t)\pi_0 + t\pi_1)_{\#}(\xi^\star)$$

is a constant speed geodesic connecting μ_0 to μ_1. Conversely any constant speed geodesic $\mu(t) : [0,1] \to \mathcal{P}_2(\mathbb{R}^d)$ connecting μ_0 to μ_1 has this represantation for some suitable plan $\xi^\star \in \Pi^\star(\mu_0,\mu_1)$.

Remark 2.3.3. *When an optimal transportation map $\mathbf{T}^\star$ exists, that pushes forward μ_0 to μ_1, the geodesic is given by*

$$t \to \mu(t) := \mu_{0\to 1}(t) = ((1-t)I + t\mathbf{T}^\star)_{\#}\,\mu_0.$$

Another useful concept that applies to the Wasserstein space is the one of the generalized geodesics.

Definition 2.3.4 (Generalized geodesics)**.** *A "generalized geodesic" joining μ_1 to μ_2 (with base μ_0) is a curve of the type*

$$\mu_{1 \underset{0}{\rightarrow} 2}(t) = (\pi_{1\rightarrow 2}(t))_{\#}\zeta, \;\; t \in [0,1],$$

where $\zeta \in \Pi(\mu_0, \mu_1, \mu_2)$, $\pi_{1\rightarrow 2}(t) = (1-t)\pi_1 + t\pi_2$ and $(\pi_{0,1})_{\#}\zeta \in \Pi^{\star}(\mu_0, \mu_1)$ and $(\pi_{0,2})_{\#}\zeta \in \Pi^{\star}(\mu_0, \mu_2)$.

Before we proceed, we are going to provide the definition of λ convex functionals along curves.

Definition 2.3.5. *Let X a metric space and a functional $\mathcal{G}\colon X \rightarrow \mathbb{R}$. We say that $\mathcal{G}$ is $\lambda-$*convex *along a curve $x\colon [0,1] \rightarrow X$ if*

$$\mathcal{G}(x(t)) \leq (1-t)\mathcal{G}(x(0)) + t\mathcal{G}(x(1)) - \frac{1}{2}\lambda t(1-t)d^2(x(0), x(1)).$$

The following definition of a geodesically convex functional is a natural extension of the previous one.

Definition 2.3.6 (**$\lambda-$geodesically convex functionals**)**.** *Let $\mathcal{G}\colon X \rightarrow \mathbb{R} \cup \{+\infty\}$. We say that $\mathcal{G}$ is λ* geodesically convex *if for any $x_0, x_1 \in \mathcal{D}(\mathcal{G})$ there exists a constant speed geodesic x with $x(0) = x_0$ and $x(1) = x_1$ such that $\mathcal{G}$ is $\lambda-$convex on x.*

In $\mathcal{P}_2(\mathbb{R}^d)$, where the notion of generalized geodesics can be defined, it is also possible to introduce the concept of convexity along generalized geodesics, which will appear very useful since in many cases it can remedy for the bad curvature properties of the Wasserstein space.

Definition 2.3.7 (**Convexity along generalized geodesics**)**.** *Let $\lambda \in \mathbb{R}$ and $\mathcal{G}\colon \mathcal{P}_2(\mathbb{R}^d) \rightarrow (-\infty, \infty]$. We say that $\mathcal{G}$ is a $\lambda-$*convex functional *if for any $\mu_0, \mu_1, \mu_2 \in \mathcal{G}$ there exists a generalized geodesic $[0,1] \in t \rightarrow \mu_{1 \underset{0}{\rightarrow} 2}(t)$ joining μ_1 to μ_2 induced by a plan ζ such that for every $t \in [0,1]$*

$$\mathcal{G}(\mu_{1\rightarrow 2}(t)) \leq (1-t)\mathcal{G}(\mu_1) + t\mathcal{G}(\mu_2) - \frac{\lambda}{2}t(1-t)\int |\mathbf{r}'' - \mathbf{r}'|^2 d\zeta(\mathbf{r}, \mathbf{r}', \mathbf{r}'').$$

2.4 Gradient flows

Gradient flows are one of the most basic type of PDE's. The usually natural interpretation on the one hand and the vast amount of numerical methods developed for finding solutions on the other make them a very popular research topic.

Let $\mathcal{G}\colon \mathbb{R}^d \to \mathbb{R}$. Let also assume that $\mathcal{G}$ is differentiable with gradient $\nabla\mathcal{G}$. We say that $\mathcal{G}$ gives rise to a gradient flow if for every $\mathbf{r}_0 \in \mathbb{R}^d$ there is a curve $\mathbf{r}\colon [0,\infty) \to \mathbb{R}^d$ such that

$$\mathbf{r}(0) = \mathbf{r}_0 \tag{2.4.1}$$

$$\frac{d\mathbf{r}(t)}{dt} = -\nabla\mathcal{G}(\mathbf{r}(t)). \tag{2.4.2}$$

When $\mathcal{G}$ is Lipschitz, then the existence of its gradient flow is guaranteed by the Picard Lindelöf theorem. The concept of gradient flows can directly be extended to the setting of Hilbert spaces, since it is possible to identify the Fréchet derivative at a point with an element of the space, by using the Riesz representation theorem.

In more general spaces without any differential structure, it is impossible to define gradient flows in the traditional sense. However if one thinks gradient flows not so much as curves $\mathbf{r}(t)$ where at every point they tend to move towards the direction of the biggest descent with speed equal to this slope, but as curves for which $\mathcal{G}(\mathbf{r}(t))$ changes as fast as it is allowed by $|\nabla\mathcal{G}(\mathbf{r}(t))|$, then it is possible to extend the concept to settings where the notion of "direction" does not exist. Indeed, the notion of gradient flows has been introduced even in the more general setting of metric spaces by De Giorgi, Degiovanni, Marino and Tosques (for more information see [10]).

To make it more specific, one has to think of

$$\begin{aligned}\frac{d(\mathcal{G}(\mathbf{r}(t)))}{dt} &= -\langle\nabla\mathcal{G}(\mathbf{r}(t)), \frac{d\mathbf{r}(t)}{dt}\rangle \geq -\left|\nabla\mathcal{G}(\mathbf{r}(t))\right| \left|\frac{d\mathbf{r}(t)}{dt}\right| \\ &\geq -\frac{1}{2}\left|\nabla\mathcal{G}(\mathbf{r}(t))\right|^2 - \frac{1}{2}\left|\frac{d\mathbf{r}(t)}{dt}\right|^2.\end{aligned}$$

The first inequality is strict unless $\frac{d\mathbf{r}(t)}{dt}$ has the same direction as $\mathcal{G}(\mathbf{r}(t))$. The second turns to equality only when $\frac{d\mathbf{r}(t)}{dt}$ and $\nabla\mathcal{G}(\mathbf{r}(t))$ have the same norm.

Only when both inequalities turn to equalities, then $|\frac{d(\mathcal{G}(\mathbf{r}(t)))}{dt}|$ achieves the biggest possible value.

Furthermore, integrating over a time interval $[t_1, t_2]$, we get formally the so called energy dissipation equality (EDE)

$$\mathcal{G}(\mathbf{r}(t_1)) = \mathcal{G}(\mathbf{r}(t_2)) + \frac{1}{2}\int_{t_1}^{t_2} \left|\frac{d\mathbf{r}(t)}{dt}\right|^2 dt + \frac{1}{2}\int_{t_1}^{t_2} |\nabla \mathcal{G}|^2(\mathbf{r}(t))dt.$$

Obviously, if we assume that $\mathbf{r}(t)$ is C^1 then the EDE holds for all gradient flows, and vice-versa (Actually the assumptions can be relaxed significantly). The advantage of the EDE over the classical gradient flow formulation is that the norm of the gradient or "maximum slope" appears instead of the gradient itself.

Before we proceed we are going to introduce the following definitions.

Definition 2.4.1 (Metric derivative). *Let (X, d) be a metric space and $x(t) : (t_0 - \epsilon, t_0 + \epsilon) \to X$ a curve in X. When it exists, we call the limit*

$$\lim_{t \to t_0} \frac{d(x(t), x(t_0))}{|t - t_0|}$$

the metric derivative of x at t_0 and we denote it $|\dot{x}|(t_0)$.

Definition 2.4.2 (Local slope). *Let (X, d) be a metric space. Let also $\mathcal{G} : X \to \mathbb{R} \cup \{\infty\}$ and $x \in X$ such that $\mathcal{G}(x) < \infty$. Then the* slope *$|\partial\mathcal{G}|(x)$ of $\mathcal{G}$ at x is*

$$|\partial\mathcal{G}|(x) := \limsup_{y \to x} \frac{(\mathcal{G}(x) - \mathcal{G}(y))^+}{d(x, y)}.$$

At this point, it is apparent that a good candidate for a gradient flow definition in setting where it is "impossible" to define the gradient is the following.

Definition 2.4.3 (Gradient flow in EDE sense). *Let (X, d) be a metric space, $\mathcal{G}\colon X \to \mathbb{R} \cup \{\infty\}$ and x_0 in $D(\mathcal{G})$. We say that $x\colon [0, \infty) \to X$ is a* gradient flow in the EDE sense *starting at x_0 if it is locally absolutely continuous curve with $x(0) = x_0$ and for every $0 \le t_1 \le t_2$*

$$\mathcal{G}(x(t_1)) = \mathcal{G}(x(t_2)) + \frac{1}{2}\int_{t_1}^{t_2} |\dot{x}|(t)^2 dt + \frac{1}{2}\int_{t_1}^{t_2} |\partial\mathcal{G}|^2(x(t))dt. \tag{2.4.3}$$

We remind the reader the notion of absolutely continuous curves.

Definition 2.4.4. *Let (X, d) be a metric space. We say that a curve $x\colon (a, b) \to X$ is p* absolutely continuous *if there exists a $g \in L^p((a, b))$ such that*

$$d(x(t_1), x(t_2)) \leq \int_{t_1}^{t_2} g(t)\, dt$$

for all $a < t_1 \leq t_2 < b$.

We will denote the set of p-absolutely continuous curves in with $\mathrm{AC}_p((a, b); X)$.

This section closes with the definition of the strong upper gradient.

Definition 2.4.5. *Let (X, d) be a metric space. A Borel function $g\colon X \to [0, +\infty]$, is a* strong upper gradient *for $\mathcal{G}$ if for every absolutely continuous curve $x \in \mathrm{AC}_1((a, b); X)$ and every $a < s \leq t < b$ we have*

$$|\mathcal{G}(x(t)) - \mathcal{G}(x(s))| \leq \int_s^t g(x(r))|\dot{x}|(r)dr.$$

In particular, if $g(x)|\dot{x}| \in L_1(a, b)$ then $\mathcal{G}(x)$ is absolutely continuous and

$$|(\mathcal{G}(x))'(t)| \leq g(x(t))|\dot{x}|(t)$$

for $\mathcal{L}_1 - a.e.$ $t \in (a, b)$.

2.4.1 Minimizing movements

After suitable notions of gradient flows in metric spaces have been established, the question of retrivability arises.

Question 2.4.6. *Is there a method that can help us retrieve the gradient flows just from the initial data ?*

In ODEs two very popular methods that work well with gradient flows are the forward and backward Euler method. Let us assume that we have a differential equation of the form

$$\frac{d\mathbf{r}(t)}{dt} = \mathbf{f}(\mathbf{r}(t), t), \ \mathbf{r}_0 = \mathbf{r}_{in}. \tag{2.4.4}$$

We define the following two iterations, where the first one corresponds to the forward and the second to the backward Euler method.

$$\mathbf{r}^{\tau}(t_{k+1}) = \mathbf{r}^{\tau}(t_k) + \tau\mathbf{f}(\mathbf{r}^{\tau}(t_k), t_k), \quad \mathbf{r}_0 = \mathbf{r}_{in}$$

and

$$\mathbf{r}^{\tau}(t_{k+1}) = \mathbf{r}^{\tau}(t_k) + \tau\mathbf{f}(\mathbf{r}^{\tau}(t_{k+1}), t_{k+1}), \quad \mathbf{r}(0) = \mathbf{r}_0, \tag{2.4.5}$$

where $t_k = \tau + t_{k-1}$.

For suitable choices of $\mathbf{f}$ (e.g. convex, Lipschitz) we get that the curve

$$\mathbf{r}^{\tau}(t) = \mathbf{r}^{\tau}(t_k) \text{ when } t \in (t_k, t_{k+1}],$$

converge to a solution of (2.4.4) when τ goes to zero.

These schemes do not only work in $\mathbb{R}^d$, but also in more general Hilbert spaces, giving a way to solve PDE's as well. Both of them have their advantages and disadvantages. Although the forward iteration seems more intuitive, the backward can work better in general setting where the functional $\mathcal{G}$ may be not properly defined for all initial data.

In the case of gradient flows, the backward Euler method takes the form

$$\mathbf{r}^{\tau}(t_{k+1}) = \mathbf{r}^{\tau}(t_k) - \tau\nabla\mathcal{G}(\mathbf{r}^{\tau}(t_{k+1})), \quad \mathbf{r}(0) = \mathbf{r}_0,$$

or

$$\frac{\mathbf{r}^{\tau}(t_{k+1}) - \mathbf{r}^{\tau}(t_k)}{\tau} + \nabla\mathcal{G}(\mathbf{r}^{\tau}(t_{k+1})) = 0, \quad \mathbf{r}(0) = \mathbf{r}_0,$$

or, if we set

$$\Phi_{\tau}(\mathbf{r}, \mathbf{r}') = \frac{|\mathbf{r} - \mathbf{r}'|^2}{2\tau} + \mathcal{G}(\mathbf{r}), \tag{2.4.6}$$

we get

$$\nabla_{\mathbf{r}}\Phi_{\tau}(\mathbf{r}^{\tau}(t_{k+1}), \mathbf{r}^{\tau}(t_k)) = 0, \quad \mathbf{r}(0) = \mathbf{r}_0 \tag{2.4.7}$$

Now if $\mathbf{r}^{\tau}(t_{k+1})$ is a minimizer of $\Phi_{\tau}(\mathbf{r}, \mathbf{r}^{\tau}(t_k))$ then (2.4.7) is satisfied, therefore one could use the following (weaker) variational iteration scheme:

$$\mathbf{r}_{k+1}^{\tau} \in \arg\min\{\frac{|\mathbf{r} - \mathbf{r}_k^{\tau}|^2}{2\tau} + \mathcal{G}(\mathbf{r})\}, \quad \mathbf{r}(0) = \mathbf{r}_0. \tag{2.4.8}$$

The above scheme has the advantage that the gradient of the functional does not appear anywhere and therefore it could be used in more general setting like the case of the metric spaces.

Definition 2.4.7 (**Discrete solution**). *Let (X, d) be a metric space, $\mathcal{G}$ be a lower semi continuous (l.s.c) functional and $x_0 \in D(\mathcal{G})$. Let also*

$$\Phi(\tau, x, y) = \Phi_\tau(x, y) = \frac{d^2(x, y)}{2\tau} + \mathcal{G}(y) \tag{2.4.9}$$

for some $\tau \in (0, \infty)$. Finally assume that a sequence $x^\tau(t_k)$ exists such that

$$x^\tau(0) = x_0, \qquad \Phi_\tau\left(x^\tau(t_{k-1}), x^\tau(t_k)\right) \leq \Phi_\tau(x^\tau(t_{k-1}), y)\ \forall y \in X.$$

The curve x_t^τ defined by

$$x^\tau(0) = x_0, \quad x^\tau(t) \ \ \textit{if} \ \ (t_{k-1}, t_k] \quad \forall k \geq 1 \tag{2.4.10}$$

is called a discrete solution

Definition 2.4.8. *Let (X, d) be a metric space, $x_0 \in X$ and $\mathcal{G} : X \to \mathbb{R}$. Let also Φ as in* (2.4.9). *We then say that a curve $x\colon [0, +\infty) \to X$ is a* generalized minimizing movement *for Φ starting at x_0 if there exists a sequence $\tau_n \to 0$ and a corresponding sequence of discrete solutions x^{τ_n} defined as in* (2.4.10) *such that*

$$\lim_{n\to\infty} \mathcal{G}(x_0^{\tau_n}) = \mathcal{G}(x_0),\ \limsup d(x_0^{\tau_n}, x_0) < \infty,\ x^{\tau_n}(t) \to x(t)\ \forall t \in [0, \infty)$$

We denote by $GMM(\Phi; x_0)$ the collection of all the generalized minimizing movements for Φ starting from x_0.

Theorem 2.4.9. *(Theorem 2.3.3 in [4]) Let (X, d) be a metric space and $\mathcal{G}$ a l.s.c functional in X. Assume further that a $\tau_* > 0$ and a $x_* \in X$ exit such that $\inf_{x\in X} \Phi_{\tau_*}(x_*; x) > -\infty$. Furthermore let $|\partial\mathcal{G}|$ be a l.s.c strong upper gradient for $\mathcal{G}$. Then every curve $x \in GMM(\Phi; x_0)$ with $x_0 \in D(\mathcal{G})$ is a curve of maximal slope* *for $\mathcal{G}$ w.r.t $|\partial\mathcal{G}|$ and in particular x satisfies the energy identity*

$$\mathcal{G}(x(0)) = \mathcal{G}(x(T)) + \frac{1}{2}\int_0^T |\dot{x}|(t)^2 dt + \frac{1}{2}\int_0^T |\partial\mathcal{G}|^2(x(t))dt, \quad \forall 0 \le T.$$

In general, there is no guarantee that a sequence of discrete solutions converges to a generalized minimizing movement. Also proving that $|\partial\mathcal{G}|$ is a l.s.c strong upper gradient is by no means a trivial matter. One assumption that guarantees both is the following.

Assumption 2.4.10 (Convexity of Φ_τ). *Let (X, d) be a metric space, $\mathcal{G}$ a functional on X and Φ_τ as in (2.4.9). Let also $\lambda \in \mathbb{R}$. For every $x_*, x_0, x_1 \in D(\mathcal{G})$ there exists a curve x with $x(0) = x_0$ and $x(1) = x_1$ such that*

$$x \to \Phi_\tau(x_*, x) \;\; \textit{is} \;\; (\tau^{-1} + \lambda) - \textit{convex on} \;\; x(t), \quad \forall\tau \;\; \textit{such that} \;\; \tau^{-1} + \lambda > 0$$

Now we have the following statements

Theorem 2.4.11. *(Corollary 2.4.10 in [4]) Suppose that $\mathcal{G}\colon X \to (-\infty, \infty]$ satisfies the convexity assumption (2.4.10)[2] some $\lambda \in \mathbb{R}$ and is lower semicontinuous. Then $|\partial\mathcal{G}|$ is a strong upper gradient for $\mathcal{G}$ and is d-lower semicontinuous.*

Theorem 2.4.12. *Let $\mathcal{G}$ such that (2.4.10) holds for some $\lambda \in \mathbb{R}$. Then the family $x^\tau(t)$ of discrete solutions with intial data x_0 converges to a function x as $\tau \to 0$, uniformly in each bounded interval $[0, T]$; in particular x is the unique element of $GMM(\Phi; x_0)$.*

For the case of the Wasserstein space it is known ([10, Lemma 4.18]) that when a functional is convex along generalized geodesics, then it satisfies the (2.4.10). Therefore by combining theorems 2.4.9, 2.4.11 and 2.4.12, we get the following statements.

Corollary 2.4.13. *Let $\mathcal{G}$ be a functional on $\mathcal{P}_2(\mathbb{R}^d)$ which is convex along generalized geodesics, and with Φ_τ as in (2.4.9) satisfying $\inf_{x\in X} \Phi_{\tau_*}(x_*; x) > -\infty$. For every $x_0 \in D(\mathcal{G})$ there exists a unique minimizing movement that it is a gradient flow of $\mathcal{G}$ in the EDE sense.*

[2] Actually the assumption can be weakened by asking that it hold only for $x_0 = x_*$.

2.5 Heat equation as a gradient flow

There are two types of functionals that are used in the sequel. The first one is potential energy functionals.

Example 2.5.1. *Let* $\Psi : \mathbb{R}^d \to (-\infty, \infty]$, *be a proper, lower semicontinuous function whose negative part has a quadratic growth, i.e.*

$$\Psi(\mathbf{r}) \geq -A - B|\mathbf{r}|^2, \ \forall \mathbf{r} \in \mathbb{R}^d \ \textit{ for some } \ A, B \in \mathbb{R}.$$

Now let the functional $\mathcal{E}$ *defined by*

$$\mathcal{E}(\mu) = \int_{\mathbb{R}^d} \Psi(\mathbf{r}) d\mu(\mathbf{r}),$$

is the potential energy functional. In cases where Ψ *is* $\lambda-$*convex, it is true that that* $\mathcal{E}$ *is* $\lambda-$ *convex along generalized geodesics.*

The second type of functionals are the internal energies.

Example 2.5.2. *Let* $G(x) : [0, \infty) \to (-\infty, \infty]$ *be a proper lower semicontinuous convex function such that*

$$G(0) = 0, \quad \liminf_{x \to 0^+} \frac{G(x)}{x^a} > -\infty, \quad a > \frac{d}{d+2}.$$

We consider the functional $\mathcal{R} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, +\infty]$ *defined by*

$$\mathcal{R}(\mu) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} G(\frac{d\mu}{d\mathbf{r}}) d\mathbf{r} & \textit{if } \mu \ll \mathcal{L} \\ \infty & \textit{otherwise} \end{cases}$$

For $G(x) = x \log x$ we get the the entropy functional (i.e $\mathcal{R}(\mu) = \mathcal{S}(\mu)$)

Both entropy $\mathcal{S}$ and $\mathcal{E}$ are convex along generalized geodesics (Propositions 9.3.2, 9.3.9. in [4]). The same holds for the free energy $\mathcal{F} = \mathcal{S} + \mathcal{E}$. Furthermore in [19] it was proved that $\inf_{\mu \in \mathcal{P}_2(\mathbb{R}^d)} \Phi_{\tau_*}(\mu_*; \mu) > -\infty$. Therefore all the assumptions of Corollary 2.4.13 are satisfied, which implies that it exists a gradient flow of the free energy functional on the Wasserstein space. Also, in [19] it is proved that the gradient flows of the free energy on the Wasserstein space, are solutions of the heat equation.

CHAPTER 3

LARGE DEVIATION PRINCIPLES AND PARTICLE MODELS

3.1 What is a large deviation?

When studying a statistical quantity (averages, moments, densities) by using a sequence of independent experiments, it is intuitively clear that after a "large" number of experiments, the empirical mean should be close to the expected value of the quantity. The above intuitive idea, in a more rigorous setting is the so called law of large numbers. Of course it remains the question of how large is large enough. When you throw a balanced dice, it is still possible (although improbable) that you will get billions of consecutive sixes and reach the unfortunate conclusion that universe has a preference to sixes. Although there is no way to know how large is large enough, the theory of large deviations can tell you how "unlucky" you have to be to make a wrong conclusion by taking into account the results from contacting many consecutive experiments.

Definition 3.1.1. *Let (X, d) be a Polish space and I a lower semicontinuous function in X. We say that a sequence X_n satisfies a Large Deviation Principle (LDP) with rate function I if, for all measurable sets $\Gamma \subset X$ we have*

$$-\inf_{x\in\Gamma^0} I(x) \leq \liminf_{n\to\infty} \frac{1}{n} \log P(X_n \in \Gamma) \leq \limsup_{n\to\infty} \frac{1}{n} \log P(X_n \in \Gamma) \leq -\inf_{x\in\overline{\Gamma}} I(x).$$

In the above definition somebody could consider X_n to be an empirical mean or an empirical measure.

The existence of such a rate function has implications both from a probabilistic and an analytic prospect. From a probabilistic point of view, having a LDP means that the probability of observing a state x different than the expected one decays exponentially with rate $I(x)$ as the number of "trials" (in our case particles) increases. From an analytical point of view, a LDP provides an "energy" functional I that characterizes the minimum effort needed to move from the expected state to another. Of course something like that makes no sense when somebody considers dice throwing, but when the object of study is particle models, the rate function often express the minimum energy that you need to insert into the system to make it deviate from its "natural" evolution and follow a different path instead.

In the above definition, when the function I has compact level sets, it is called a good rate function. Most examples in literature are with good rate functions, but it is possible to encounter different types as well. Also there are some weaker notions of LDP. For a nice exposition the reader can go to [12]

One of the most simple and at the same useful theorems in the theory of large deviations, is the contraction principle.

Theorem 3.1.2. *(Contraction principle) Let $(X, d), (X', d')$ be two Polish spaces and $f : X \to X'$ a continuous function. Consider a good rate function $I : X \to [0, \infty]$.*

- *for each $x' \in X'$, define*

$$I'(x') = \inf\{I(x) : x \in X, x' = f(x)\}.$$

  Then I' is a good rate function on X'.

- *If I is the rate function for the LDP associated with a sequence of variables $\{X_n\}$ on X, then I' is the rate function for the LDP associated with the sequence $f(X_n)$ on X'*

The contraction principle has an intuitive interpretation as well. If someone thinks of X as a space of possible ways leading to outcomes that are elements in

X', then an outcome is as improbable to occur as the most probable way that it can happen.

3.2 Particle model.

The sequence of models studied, is the one of independent Brownian particles starting their movement from fixed (deterministic) positions. For every $N \in \mathbb{N}$, let $\mathbf{r}_1, \ldots, \mathbf{r}_N$ be the solutions to the Itō stochastic equations

$$d\mathbf{r}_i(t) = \sqrt{2}\, d\mathbf{w}_i(t), \quad \mathbf{r}_i(0) = \mathbf{r}_{i,0}, \qquad i = 1, \cdots, N$$

where $\mathbf{w}_1, \ldots, \mathbf{w}_N$ are independent Wiener processes.

With $L_N(t) = \frac{1}{N}\sum_{i=1}^{N} \delta_{\mathbf{r}_i(t)}$ we define the empirical measure. If we assume that $L_N(0) \rightharpoonup \mu_0$, then we have that $L_N(t)$ converges narrowly to the solution of

$$\partial_t \mu(t) = \Delta\mu(t), \quad \mu(0) = \mu_0.$$

It is a classical result by know that, L_n also satisfies a large deviation principle [11],[20]. The rate function that we derive from the large deviation is

$$J_\tau(\mu|\mu_0) = \inf\Big\{\mathcal{H}(\xi\|\mu_0 P_\tau) : \xi \in \Pi(\mu_0, \mu)\Big\},$$

where $\mathcal{H}$ is the relative entropy and P_τ is the transition kernel for the heat equation.

We remind the reader that the relative entropy is defined as follows.

$$\mathcal{H}(\xi'\|\xi) = \begin{cases} \displaystyle\int \frac{d\xi'}{d\xi}\log(\frac{d\xi'}{d\xi})d\xi & \text{when } \ \xi' \ll \xi \\ \infty & \text{otherwise.} \end{cases} \tag{3.2.1}$$

Now that the intuitive idea behind the LDPs has been explained it is possible to give a reasoning behind the result in [20] (i.e that $\tau J_\tau(\cdot|\mu_0)$ Gamma-converges to the Wasserstein metric $W_2(\cdot, \mu_0)$). For very small values of τ, where diffusion does not have enough time to act, the energy needed to go to a different state depends only on the dissipation mechanism, which is the Wasserstein distance in this case. To answer the question why is the Wasserstein distance the dissipation mechanism, someone has to look at the transition kernel for the heat equation.

CHAPTER 4

MAIN RESULT: PART 1. THE CASE OF THE GAUSSIAN SUBMANIFOLD.

4.1 The central statement

In this chapter the main conjecture, is being proved for the special case where both the set of initial data and the domain of J is the set of all Gaussians

$$\mathcal{G}(\mathbb{R}) = \left\{ \mu = \mathcal{N}(m, \sigma) | \frac{d\mathcal{N}(m, \sigma)}{dx} = \exp\left(\frac{-(x-m)^2}{2\sigma^2} \right) \right\}.$$

Let $\mu_0 \in \mathcal{G}(\mathbb{R})$ and $\Psi \equiv 0$. Let also J_τ the rate function that we obtain in the previous chapter via the LDP for the empirical measure. We have

$$J_\tau(\cdot\,|\mu_0) - \frac{W_2^2(\mu_0, \cdot\,)}{4\tau} \to \frac{1}{2}\mathcal{S}(\,\cdot\,) - \frac{1}{2}\mathcal{S}(\mu_0), \text{ in } \quad \mathcal{G}(\mathbb{R}).$$

At this point, it is normal for the reader, to wonder if it makes any sense to restrict the problem to the case of the Gaussians and also what could be the reason. One of the main reasons to work with Gaussians, is that it is quite convenient. Many quantities have already been calculated explicitly (Wasserstein distance, entropy) and it seems more than likely, that the minimizers in the definition of

[0]This chapter is joint work with Nicolas Dirr and Johannes Zimmer ([13])

J_τ as well as its values could be calculated explicitly as well. Furthermore, even from a theoretical/philosophical point of view, the space $\mathcal{G}(\mathbb{R})$ is a quite natural restriction for simple diffusion equation. Not only is the set $\mathcal{G}(\mathbb{R})$ closed with respect to diffusion, but also the iteration process $\arg\min\{\frac{W_2^2(\mu_0,\mu)}{2\tau}+\mathcal{S}(\mu)\}$ always generates measures inside $\mathcal{G}(\mathbb{R})$, when $\mu_0 \in \mathcal{G}$.

Theorem 4.1.1. *Let $\mathcal{N}(m_0, \sigma_0^2)$ be a normal distribution. Then for the rate functional J_τ it holds that*

$$J_\tau(\cdot;\mathcal{N}(m_0,\sigma_0^2)) - \frac{W_2^2(\cdot,\mathcal{N}(m_0,\sigma_0^2))}{4\tau} \to \frac{1}{2}\mathcal{S}(\cdot) - \frac{1}{2}\mathcal{S}(\mathcal{N}(m_0,\sigma_0^2)),$$

locally uniform and in the sense of Γ-convergence with respect to the weak topology, on the submanifold of the Gaussians.

The proof of the central statement is as follows. First we explicitly calculate the minimizer of $\mathcal{H}(\cdot\|\xi_{0\to\tau})$ for every τ, μ and μ_0. Then we estimate the difference with the functional $\frac{1}{2}\Phi_\tau(\mu_0;\mu)$; the statement then follows easily from the topological structure of the Gaussian submanifold.

4.2 Minimising the relative entropy over a bivariate

Theorem 4.2.1. *Let ξ be absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$. Furthermore let Ξ' be the set of all Borel measures ξ' in $\mathbb{R}^d$ satisfying*

$$\int q_i(\mathbf{r}) \cdot \xi'(\mathbf{r})d\mathbf{r} = a_i\ ,\ i \in \{1,2,...N\}, \tag{4.2.1}$$

where $q_i(\mathbf{r})\colon \mathbb{R}^d \to \mathbb{R}$ are given functions and $a_i \in \mathbb{R}$. Let us finally assume that there is a measure $\xi^\star \in \Xi'$ that has a density of the form

$$\xi^\star(\mathbf{r}) = \xi(\mathbf{r}) \exp\left(\sum_{i=1}^{n} \lambda_i q_i(\mathbf{r})\right)$$

for some $\lambda_i \in \mathbb{R}$. Then $\xi^\star$ is the unique minimiser of $\mathcal{H}(\xi'\|\xi)$ over all $\xi' \in \Xi'$.

Proof. We can assume that $\xi' \ll \xi$ since otherwise we trivially have that $H(\xi' \| \xi) = \infty$. Due to the form of $\xi^\star$ it is also true that $\xi' \ll \xi^\star$. Now

$$\begin{aligned}
\mathcal{H}(\xi' \| \xi) &= \int \xi'(\mathbf{r}) \log \frac{\xi'(\mathbf{r})}{\xi(\mathbf{r})} d\mathbf{r} = \int \xi'(\mathbf{r}) \log \frac{\xi'(\mathbf{r})}{\xi^\star(\mathbf{r})} d\mathbf{r} + \int \xi'(\mathbf{r}) \log \frac{\xi^\star(\mathbf{r})}{\xi(\mathbf{r})} d\mathbf{r} \\
&= H(\xi' \| \xi^\star) + \int \xi'(\mathbf{r}) \log \frac{\xi^\star(\mathbf{r})}{\xi(\mathbf{r})} d\mathbf{r} \geq \int \xi'(\mathbf{r}) \log e^{\sum_{i=1}^n \lambda_i q_i(\mathbf{r})} d\mathbf{r} \\
&= \sum_{i=1}^N \lambda_i \int \xi'(\mathbf{r}) \cdot q_i(\mathbf{r}) d\mathbf{r} = \sum_{i=1}^N \lambda_i a_i = \sum_{i=1}^N \lambda_i \int \xi^\star(\mathbf{r}) \cdot q_i(\mathbf{r}) d\mathbf{r} \\
&= \int \xi^\star(\mathbf{r}) \log e^{\sum_{i=1}^N \lambda_i q_i(\mathbf{r})} d\mathbf{r} = \int \xi^\star(\mathbf{r}) \log \frac{\xi^\star(\mathbf{r})}{\xi(\mathbf{r})} d\mathbf{r} = \mathcal{H}(\xi^\star \| \xi),
\end{aligned}$$

where the first inequality is an equality if and only if $\xi^\star = \xi'$, by the properties of the relative entropy functional. So $\xi^\star$ is the unique minimizer. $\square$

Below we will calculate the relative entropy of one bivariate with respect to another. We recall that a measure is a *bivariate* with marginals $\mu_1 = \mathcal{N}(m_1, \sigma_1^2)$, $\mu_2 = \mathcal{N}(m_2, \sigma_2^2)$ and "correlation" θ if it has Lebesgue density

$$\begin{aligned}
\nu(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\theta^2}} \exp\Bigg(- \frac{1}{2(1-\theta^2)} \bigg[\frac{(x-m_1)^2}{\sigma_1^2} + \frac{(y-m_2)^2}{\sigma_2} \\
- \frac{2\theta(x-m_1)(y-m_2)}{\sigma_1\sigma_2} \bigg] \Bigg);
\end{aligned}$$

we then write $\mathcal{N}(m_1, \sigma_1^2, m_2, \sigma_2^2, \theta)$ for this measure. To every such bivariate we associate

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \theta\sigma_1\sigma_2 \\ \theta\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \ \boldsymbol{m} = \begin{pmatrix} m_1 \\ m_2 \end{pmatrix},$$

to write more succinctly, with $\mathbf{r} = (x_1, x_2)$,

$$\nu(\mathbf{r}) = \frac{1}{2\pi\sqrt{\det|\Sigma|}} \exp\left(-\frac{1}{2} (\mathbf{r} - \boldsymbol{m})^T \Sigma^{-1} (\mathbf{r} - \boldsymbol{m}) \right)$$

where the first matrix is known as the *covariance matrix* of the bivariate. For two bivariates ξ, ξ', the relative entropy can be expressed [23] as

$$\mathcal{H}(\xi^{'} \| \xi) = \frac{1}{2} \left[\operatorname{tr}(\Sigma^{-1}\Sigma^{'}) + (\boldsymbol{m}^{'} - \boldsymbol{m})^T \Sigma^{-1} (\boldsymbol{m}^{'} - \boldsymbol{m}) - \log \left(\frac{\det \Sigma^{'}}{\det \Sigma} \right) - 2 \right]. \tag{4.2.2}$$

Theorem 4.2.2. *Let ξ a bivariate with marginals $\mathcal{N}(m_1, \sigma_1^2)$ and $\mathcal{N}(m_2, \sigma_2^2)$ and correlation θ. Let $\Pi\big(\mathcal{N}(m_1^\star, \sigma_1^{\star 2}), \mathcal{N}(m_2^\star, \sigma_2^{\star 2})\big)$ be the set of all Borel measures with marginals $\mathcal{N}(m_i^\star, \sigma_i^{\star 2})$, for $i = 1, 2$. Then the relative entropy functional $\mathcal{H}(\cdot \| \xi)$ has a unique minimizer in $\Pi\big(\mathcal{N}(m_1^\star, \sigma_1^\star), \mathcal{N}(m_2^\star, \sigma_2^\star)\big)$, namely the bivariate $\xi^\star$ with correlation $\theta^\star$ given by*

$$\frac{\theta^\star}{1 - \theta^{\star 2}} = \frac{\theta}{1 - \theta^2} \cdot \frac{\sigma_1^\star \sigma_2^\star}{\sigma_1 \sigma_2}. \tag{4.2.3}$$

Proof. We actually prove a stronger statement, namely that $\xi^\star$ is the minimizer of $\mathcal{H}(\cdot \| \xi)$ over Ξ', which is defined as the set of all ξ satisfying

$$\int \xi'(x, y) dx dy = 1,$$

$$\int x \xi'(x, y) dx dy = m_1^\star, \qquad \int y \xi'(x, y) dx dy = m_2^\star,$$

$$\int x^2 \xi'(x, y) dx dy = \sigma_1^{\star 2} + m_1^{\star 2}, \qquad \int y^2 \xi'(x, y) dx dy = \sigma_2^{\star 2} + m_2^{\star 2}.$$

Since $\xi^\star \in \Pi\big(\mathcal{N}(m_1^\star, \sigma_1^\star), \mathcal{N}(m_2^\star, \sigma_2^\star)\big)$ and $\Pi\big(\mathcal{N}(m_1^\star, \sigma_1^\star), \mathcal{N}(m_2^\star, \sigma_2^\star)\big) \subset \Xi'$, it follows that $\xi^\star$ is also a minimizer in $\Pi\big(\mathcal{N}(m_1^\star, \sigma_1^\star), \mathcal{N}(m_2^\star, \sigma_2^\star)\big)$.

The densities $\xi^\star$ and ξ of $\xi^\star$ and ξ satisfy

$$\frac{\xi^\star}{\xi} = \frac{\frac{1}{2\pi\sigma_1^\star\sigma_2^\star\sqrt{1-\theta^{\star 2}}}\exp\left(-\frac{1}{2(1-\theta^{\star 2})}\left[\frac{(x-m_1^\star)^2}{\sigma_1^{\star 2}}+\frac{(y-m_2^\star)^2}{\sigma_2^{\star 2}}-\frac{2\theta^\star(x-m_1^\star)(y-m_2^\star)}{\sigma_1^\star\sigma_2^\star}\right]\right)}{\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\theta^2}}\exp\left(-\frac{1}{2(1-\theta^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}+\frac{(y-m_2)^2}{\sigma_2^2}-\frac{2\theta(x-m_1)(y-m_2)}{\sigma_1\sigma_2}\right]\right)} =$$
$$\frac{2\pi\sigma_1\sigma_2\sqrt{1-\theta^2}}{2\pi\sigma_1^\star\sigma_2^\star\sqrt{1-\theta^{\star 2}}}\exp\left(\frac{1}{2(1-\theta^2)}\left[\frac{(x-m_1)^2}{\sigma_1^2}+\frac{(y-m_2)^2}{\sigma_2^2}-\frac{2\theta(x-m_1)(y-m_2)}{\sigma_1\sigma_2}\right]\right.$$
$$\left.-\frac{1}{2(1-\theta^{\star 2})}\left[\frac{(x-m_1^\star)^2}{\sigma_1^{\star 2}}+\frac{(y-m_2^\star)^2}{\sigma_2^{\star 2}}-\frac{2\theta^\star(x-m_1^\star)(y-m_2^\star)}{\sigma_1^\star\sigma_2^\star}\right]\right) =$$
$$\frac{\sigma_1\sigma_2\sqrt{1-\theta^2}}{\sigma_1^\star\sigma_2^\star\sqrt{1-\theta^{\star 2}}}\exp\left(\left[\frac{1}{2\sigma_1^2(1-\theta^2)}-\frac{1}{2\sigma_1^{\star 2}(1-\theta^{\star 2})}\right]x^2+\right.$$
$$\left[\frac{1}{2\sigma_2^2(1-\theta^2)}-\frac{1}{2\sigma_2^{\star 2}(1-\theta^{\star 2})}\right]y^2+\left[\frac{\theta^\star}{(1-\theta^{\star 2})\sigma_1^\star\sigma_2^\star}-\frac{\theta}{(1-\theta^2)\sigma_1\sigma_2}\right]xy+$$
$$\left[\frac{m_1}{(1-\theta^2)\sigma_1^2}-\frac{\theta m_2}{(1-\theta^2)\sigma_1\sigma_2}+\frac{m_1^\star}{(1-\theta^{\star 2})\sigma_1^{\star 2}}-\frac{\theta^\star m_2^\star}{(1-\theta^{\star 2})\sigma_1^\star\sigma_2^\star}\right]x-$$
$$\left[\frac{m_1}{(1-\theta^2)\sigma_1^2}-\frac{\theta m_2}{(1-\theta^2)\sigma_1\sigma_2}+\frac{m_1^\star}{(1-\theta^{\star 2})\sigma_1^{\star 2}}-\frac{\theta^\star m_2^\star}{(1-\theta^{\star 2})\sigma_1^\star\sigma_2^\star}\right]y+$$
$$\left.\left[\frac{\theta m_1 m_2}{\sigma_1\sigma_2(1-\theta^2)}-\frac{\theta^\star m_1^\star m_2^\star}{\sigma_1^\star\sigma_2^\star(1-\theta^{\star 2})}\right]\right)\overset{(4.2.3)}{=}$$
$$\exp(Ax^2+By^2+Cx+Dy+E),$$

for suitable constants A, B, C, D, E. Thus $\xi^\star$ has a density of the form

$$\xi^\star(\mathbf{r}) = \xi(\mathbf{r})\exp\left(\sum_{i=1}^n \lambda_i q_i(\mathbf{r})\right).$$

Therefore by Theorem 4.2.1 $\xi^\star$ minimizes the functional $\mathcal{H}(\cdot\|\xi)$ over

$$\Pi\left(\mathcal{N}(m_1^\star,\sigma_1^\star),\mathcal{N}(m_2^\star,\sigma_2^\star)\right)$$

. $\square$

4.3 Asymptotic behavior of the rate functional

Theorem 4.3.1. *Let $\mathcal{N}(m_0, \sigma_0^2)$ and $\mathcal{N}(m, \sigma^2)$ be two normal distributions. Then for the rate functional we have*

$$J_\tau(\mathcal{N}(m, \sigma^2)|\mathcal{N}(m_0, \sigma_0^2)) = \frac{1}{2}\left[\frac{(\sigma - \sigma_0)^2 + (m - m_0)^2}{2\tau} + \frac{1 - \sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2} + 1} - \frac{\tau}{\sigma_0\sigma}}{\frac{\tau}{\sigma\sigma_0}} - \log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2} + 1} - \frac{\tau}{\sigma_0\sigma}\right) - \log\frac{\sigma}{\sigma_0} - 1\right]. \tag{4.3.1}$$

Proof. We have

$$J_\tau(\mathcal{N}(m, \sigma^2); \mathcal{N}(m_0, \sigma_0^2)) = \inf_{\xi \in \Pi(\mathcal{N}(m_0, \sigma_0^2), \mathcal{N}(m, \sigma^2))} H(\xi \| \xi_{0\to\tau}),$$

where $\xi_{0\to\tau} = \mathcal{N}(m_0, \sigma_0^2)P_\tau$. The probability distribution of $\xi_{0\to\tau}$ is given by

$$\xi_{0\to\tau} = \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(\frac{-(x - m_0)^2}{2\sigma_0^2}\right) \cdot \frac{1}{\sqrt{4\pi\tau}} \exp\left(\frac{-(x - y)^2}{4\tau}\right).$$

It is easy to see that $\xi_{0\to\tau}$ is then the Bivariate

$$\mathcal{N}(m_0, \sigma_0^2, m_0, \sigma_\tau^2, \theta) \quad \text{where} \quad \sigma_\tau^2 = \sigma_0^2 + 2\tau \quad \text{and} \quad \theta = \frac{\sigma_0}{\sigma_\tau}; \tag{4.3.2}$$

for the corresponding matrices

$$\Sigma_{0\to\tau} = \begin{pmatrix} \sigma_0^2 & \theta\sigma_0\sigma_\tau \\ \theta\sigma_0\sigma_\tau & \sigma_\tau^2 \end{pmatrix}, \ \boldsymbol{m}_{0\to\tau} = \begin{pmatrix} m_0 \\ m_0 \end{pmatrix},$$

we have trivially

$$\det(\Sigma_{0\to\tau}) = \sigma_\tau^2\sigma_0^2 - \theta^2\sigma_0^2\sigma_\tau^2 = (1 - \theta^2)\sigma_0^2\sigma_\tau^2 = 2\tau\sigma_0^2$$

and

$$\Sigma_{0\to\tau}^{-1} = \frac{1}{2\tau\sigma_0^2}\begin{pmatrix} \sigma_\tau^2 & -\theta\sigma_0\sigma_\tau \\ -\theta\sigma_0\sigma_\tau & \sigma_0^2 \end{pmatrix}.$$

By Theorem 4.2.2, it follows that the unique minimizer for $\mathcal{H}(\cdot\|\xi_{0\to\tau})$ among the measures in $\Pi\left(\mathcal{N}(m_0,\sigma_0),\mathcal{N}(m,\sigma)\right)$ is the bivariate $\xi^\star=\mathcal{N}(m_0,\sigma_0^2,m,\sigma^2,\theta^\star)$, with corresponding matrix

$$\Sigma^\star=\begin{pmatrix}\sigma_0^2 & \theta^\star\sigma_0\sigma\\ \theta^\star\sigma_0\sigma & \sigma^2\end{pmatrix},\ \boldsymbol{m}^\star=\begin{pmatrix}m_0\\ m\end{pmatrix},$$

where $\theta^\star$ satisfies

$$\frac{\theta^\star}{1-\theta^{\star 2}}=\frac{\theta}{1-\theta^2}\cdot\frac{\sigma_0\sigma}{\sigma_0\sigma_\tau}=\frac{\sigma_0\sigma_\tau}{2\tau}\cdot\frac{\sigma_0\sigma}{\sigma_0\sigma_\tau}=\frac{\sigma_0\sigma}{2\tau}.$$

By solving the above quadratic equation, we find

$$\theta^\star=\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}-\frac{\tau}{\sigma_0\sigma}. \tag{4.3.3}$$

So for the rate functional as in (4.2.2),

$$\begin{aligned}&\mathcal{H}(\xi^\star\|\xi_{0\to\tau})\\ &=\frac{1}{2}\left[\operatorname{tr}\left(\Sigma_{0\to\tau}^{-1}\Sigma^\star\right)+(\boldsymbol{m}^\star-\boldsymbol{m})^T\Sigma_{0\to h}^{-1}(\boldsymbol{m}^\star-\boldsymbol{m})-\log\left(\frac{\det\Sigma^\star}{\det\Sigma_{0\to\tau}}\right)-2\right],\end{aligned}$$

we have

$$\begin{aligned}\operatorname{tr}\left(\Sigma_{0\to\tau}^{-1}\Sigma^\star\right)&=\operatorname{tr}\left(\frac{1}{2\tau\sigma_0^2}\begin{pmatrix}\sigma_\tau^2 & -\theta\sigma_0\sigma_\tau\\ -\theta\sigma_0\sigma_\tau & \sigma_0^2\end{pmatrix}\cdot\begin{pmatrix}\sigma_0^2 & \theta^\star\sigma_0\sigma\\ \theta^\star\sigma_0\sigma & \sigma^2\end{pmatrix}\right)\\ &=\frac{\sigma_\tau^2\sigma_0^2-2\theta\theta^\star\sigma_0^2\sigma\sigma_\tau+\sigma_0^2\sigma^2}{2\tau\sigma_0^2}=\frac{\sigma_\tau^2-2\theta\theta^\star\sigma\sigma_\tau+\sigma^2}{2\tau}\\ &\overset{(4.3.2)}{=}\frac{\sigma_0^2+2\tau-2\theta\theta^\star\sigma\sigma_\tau+\sigma^2}{2\tau}\\ &=\frac{(\sigma-\sigma_0)^2}{2\tau}+\frac{2\sigma\sigma_0-2\theta\theta^\star\sigma\sigma_\tau}{2\tau}+1\\ &\underset{(4.3.3)}{\overset{(4.3.2)}{=}}\frac{(\sigma-\sigma_0)^2}{2\tau}+\frac{1+\frac{\tau}{\sigma_0\sigma}-\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}}{\frac{\tau}{\sigma\sigma_0}}+1.\end{aligned} \tag{4.3.4}$$

Also,

$$
\begin{aligned}
&(\boldsymbol{m}^\star - \boldsymbol{m})^T \Sigma_{0\to\tau}^{-1} (\boldsymbol{m}^\star - \boldsymbol{m}) && (4.3.5)\\
=&\begin{pmatrix} 0 \\ m - m_0 \end{pmatrix}^T \frac{1}{2\tau \sigma_0^2} \begin{pmatrix} \sigma_\tau^2 & -\theta\sigma_0\sigma_\tau \\ -\theta\sigma_0\sigma_\tau & \sigma_0^2 \end{pmatrix} \begin{pmatrix} 0 \\ m - m_0 \end{pmatrix}\\
=&\frac{(m-m_0)^2\sigma_0^2}{2\tau\sigma_0^2} = \frac{(m-m_0)^2}{2\tau}. && (4.3.6)
\end{aligned}
$$

Finally,

$$
\begin{aligned}
\log\left(\frac{\det \Sigma^\star}{\det \Sigma_{0\to\tau}}\right) &= \log\left(\frac{(1-\theta^{\star 2})\sigma_0^2\sigma^2}{(1-\theta^2)\sigma_0^2\sigma_\tau^2}\right) \overset{(4.2.3)}{=} \log \frac{\theta^\star \sigma_0^2\sigma^2\sigma_0\sigma_\tau}{\theta\sigma_0^2\sigma_\tau^2\sigma_0\sigma} = \log\frac{\theta^\star}{\theta} + \log\frac{\sigma}{\sigma_\tau}\\
&\underset{(4.3.3)}{\overset{(4.3.2)}{=}} \log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1} - \frac{\tau}{\sigma_0\sigma}\right) - \log\frac{\sigma_0}{\sigma_\tau} + \log\frac{\sigma}{\sigma_\tau}\\
&= \log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1} - \frac{\tau}{\sigma_0\sigma}\right) + \log\frac{\sigma}{\sigma_0}. && (4.3.7)
\end{aligned}
$$

So by (4.3.4), (4.3.6) and (4.3.7), we get

$$
\begin{aligned}
2\mathcal{H}(\xi^\star \| \xi_{0\to\tau}) = \Bigg[&\frac{(\sigma-\sigma_0)^2 + (m-m_0)^2}{2\tau} + \frac{1 - \sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1} - \frac{\tau}{\sigma_0\sigma}}{\frac{\tau}{\sigma\sigma_0}}\\
&- \log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1} - \frac{\tau}{\sigma_0\sigma}\right) - \log\frac{\sigma}{\sigma_0} - 1\Bigg]
\end{aligned}
$$

as claimed. □

We continue with the proof of the central statement, Theorem 4.1.1. We start with two observations. First, for Gaussian measures, the weak topology is equivalent to the one induced by the Wasserstein metric [29, Theorem 7.12]. Second, by [25], for two Gaussians $\mathcal{N}(m_0, \sigma_0^2)$ and $\mathcal{N}(m, \sigma^2)$, it holds that

$$
W^2(\mathcal{N}(m_0,\sigma_0^2), \mathcal{N}(m,\sigma^2)) = (m-m_0)^2 + (\sigma-\sigma_0)^2. \tag{4.3.8}
$$

And so

$$
\begin{aligned}
&\Bigg[J_\tau(\mathcal{N}(m,\sigma^2)|\mathcal{N}(m_0,\sigma_0^2)) - \frac{W_2^2(\mathcal{N}(m,\sigma^2),\mathcal{N}(m_0,\sigma_0^2))}{4\tau} \qquad (4.3.9)\\
&- \frac{1}{2}\mathcal{S}(\mathcal{N}(m,\sigma^2)) + \frac{1}{2}\mathcal{S}(\mathcal{N}(m_0,\sigma_0^2))\Bigg]\\
&\le \frac{1}{2}\left|\frac{1+\frac{\tau}{\sigma_0\sigma}-\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}}{\frac{\tau}{\sigma\sigma_0}} - \log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}-\frac{\tau}{\sigma_0\sigma}\right)-1\right|\\
&\le \left|\frac{1+\frac{\tau}{\sigma_0\sigma}-\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}}{\frac{2\tau}{\sigma\sigma_0}} - \frac{1}{2}\right| + \frac{1}{2}\left|\log\left(\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}-\frac{\tau}{\sigma_0\sigma}\right)\right|\\
&\le \left|\frac{1-\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}}{\frac{2\tau}{\sigma\sigma_0}}\right| + \frac{1}{2}\left|\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}-\frac{\tau}{\sigma_0\sigma}-1\right|\\
&\le \left|\frac{1-\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}}{\frac{2\tau}{\sigma\sigma_0}}\right| + \left|\frac{1}{2}\frac{2\frac{\tau}{\sigma\sigma_0}}{\sqrt{\frac{\tau^2}{\sigma_0^2\sigma^2}+1}+\frac{\tau}{\sigma_0\sigma}+1}\right|\\
&\le \frac{o(\tau)}{\sigma}. \qquad (4.3.10)
\end{aligned}
$$

Now for fixed $\delta > 0$, we have uniform convergence for $\sigma' > \delta$. For the Γ-convergence, we get the lower bound part by the local uniform convergence and the lower semi-continuity of the limit. As a recovery sequence we can choose the fixed sequence $\mu_\tau = \mathcal{N}(m,\sigma^2)$. □

CHAPTER 5

THE DIFFERENTIAL STRUCTURE OF THE WASSERSTEIN SPACE

5.1 The tangent cone

One of the most basic equations in the theory of fluid dynamics is the continuity equation. In a formal way, it is given by

$$\partial_t \mu(t) + \operatorname{div}(\mu(t)\, \mathbf{v}(t)) = 0 \tag{5.1.1}$$

and it is satisfied in systems where the rate at which mass enters is equal to the rate that exits[1].

Usually the above equality is meant in a distributional sense, i.e.

$$\int_0^T \int_{\mathbb{R}^d} (\partial_t f(t,\mathbf{r}) + < \mathbf{v}(t,\mathbf{r}), \nabla_{\mathbf{r}} f(t,\mathbf{r}) >)\, \mu(t)(d\mathbf{r})dt,\ \forall f \in D([0,T]\times\mathbb{R}^d).$$

In the above definition $\mathcal{D}(A)$, where A is a set in $[0,T]\times\mathbb{R}^d$ (or sometimes just in $\mathbb{R}^d$), is the set of all infinitely many times differentiable functions with compact support (a.k.a test functions). Also with $\mathcal{D}'(A)$ it will be denoted the

[1]For visualisation purposes, one can consider that $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ and $v(t) \in L^2(\mu(t);\mathbb{R}^d)$, which is actually our case, however this equation can make sense in more general settings.

dual of $\mathcal{D}(A)$[1] (a.k.a the set of all distributions in A). When A is obvious from the context, it will be omitted.

For all t such that $\partial_t \mu(t)$ is well defined (i.e. makes sense pointwise) and it is given by a function $\partial_t \mu(t, \mathbf{r})$, the quantity $\partial_t \mu(t, \mathbf{r})$ characterizes the rate that the mass changes at point $\mathbf{r}$ and the vector $\mathbf{v}(t, \mathbf{r})$ describes the direction and the speed at which the mass at point $\mathbf{r}$ tends to move at time t.

At this point it is natural to ask when does a curve $\mu(t)$ satisfies the continuity equation for some vector field $\mathbf{v}(t)$ and conversely, when does a vector field $\mathbf{v}(t)$ gives rise to a curve $\mu(t)$ that satisfies the continuity equation with respect to that vector field.

Although there are many relevant results in the literature, we will restrict ourselves in the case of the Wasserstein absolutely continues curves, where the following holds.

Theorem 5.1.1. *[3, Th. 8.3.1] Let I be an open interval in $\mathbb{R}$ and $\mu(t) : I \to \mathcal{P}_2(\mathbb{R}^d)$ be an absolutely continuous curve with respect to the Wasserstein topology. Also let assume that the metric derivative $|\dot{\mu}|_t$ is in $L^1(I)$. Then there exists a vector field $\mathbf{v}(t, \mathbf{r}) : I \times \mathbb{R}^d \to \mathbb{R}^d$ such that*

$$\mathbf{v}(t) \in L^2(\mu(t); \mathbb{R}^d), \qquad \|\mathbf{v}(t)\|_{L^2(\mu(t);\mathbb{R}^d)} \leq |\dot{\mu}|(t) \quad \textit{for a.e } t \in I \tag{5.1.2}$$

and the continuity equation $\partial_t \mu(t) + \operatorname{div}(\mu(t)\, \mathbf{v}(t)) = 0$, holds in the distributional sense. Furthermore for every $t \in I$, we have that $\mathbf{v}(t)$ belongs in the $L^2(\mu(t); \mathbb{R}^d)$ closure of the subspace generated by ∇f with $f \in \mathcal{D}$.

Conversely let $\mu(t) : I \to \mathcal{P}(\mathbb{R}^d)$ be a narrowly continuous curve that satisfies the continuity equation for some vector field $\mathbf{v}(t)$ with

$$\int_I \|\mathbf{v}(t)\|^2_{L^2(\mu(t))} dt < \infty. \tag{5.1.3}$$

Let also assume that $\mu(t_0) \in \mathcal{P}_2(\mathbb{R}^d)$ for some $t_0 \in I$, then $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ for all $t \in I$ and $\mu(t)$ is absolutely continuous in the Wasserstein sense with $|\dot{\mu}|(t) \leq \|\mathbf{v}(t)\|_{L^2(\mu(t);\mathbb{R}^d)}$

[1]For more details regarding the actual topology on $\mathcal{D}, \mathcal{D}'$ the reader can look in [24].

Remark 5.1.2. *We point out that the hypothesis in [3, Th. 8.3.1] requires a priori that the curve $\mu(t)$ lies in $\mathcal{P}_2(\mathbb{R}^d)$, but the proof actually shows that $\mu(t_0) \in \mathcal{P}_2(\mathbb{R}^d)$ for some $t_0 \in I$, implies the whole curve to be in $\mathcal{P}_2(\mathbb{R}^d)$ (and it is absolutely continuous in the Wasserstein sense).*

As Theorem 5.1.1 shows, from all vector fields $\mathbf{v}(t)$ that gives rise to a specific curve $\mu(t)$ there is at least one that minimizes the integral

$$\int_I \|\mathbf{v}(t)\|^2_{L^2(\mu(t);\mathbb{R}^d)} dt < \infty.$$

Moreover, the ones that minimize it are the ones for which

$$\mathbf{v}(t) \in \overline{\{\nabla f : f \in D\}}^{L^2(\mu(t);\mathbb{R}^d)},$$

for a.e. $t \in I$

Motivated by the above one, can define a tangent cone $Tan_\mu \mathcal{P}_2(\mathbb{R}^d)$ for every point μ in $\mathcal{P}_2(\mathbb{R}^d)$. Before we proceed we are going to introduce three Hilbert spaces (depending on $\mu \in \mathcal{P}_2(\mathbb{R}^d)$) along with a very natural isometry between them. Let, $\|f\|^2_{1,\mu} = \int_{\mathbb{R}^d} |\nabla f|^2 d\mu,\ \forall f \in D$. The first space is $\dot{H}_\mu$ and it is the completion of $\mathcal{D}$ under $\|\cdot\|^2_{1,\mu}$.[2]

Another space is $L^2_{\mu,\nabla}(\mathbb{R}^d) = \overline{\{\nabla f : f \in D\}}^{L^2_\mu(\mathbb{R}^d)}$.

The last space is the dual of $\dot{H}_\mu$, i.e.

$$\dot{H}^{-1}_\mu = \{\text{equivalence class of } s \in D' : \|s\|_{-1,\mu} < \infty\},$$

where

$$\|s\|^2_{-1,\mu} := \sup_{f \in \mathcal{D}} \left\{ \langle s, f \rangle - \frac{1}{2} \int_{\mathbb{R}^d} |\nabla f|^2 d\mu \right\}. \tag{5.1.4}$$

The connection (isometry) between them is given by [16, Lem. D.34], where we get that for every $s \in \dot{H}^{-1}_\mu$ it exists a $f \in \dot{H}_\mu$ with $v = \nabla f \in L^2_{\mu,\nabla}(\mathbb{R}^d)$ such that

$$s = -\nabla(\mu\, \mathbf{v}) = -\nabla(\mu\, \nabla f) \qquad \text{in distributional sense.} \tag{5.1.5}$$

[2]Here and in the sequel, functions in $\mathcal{D}$ with gradients equal $\mu-$a.e. will be considered the same element in $\dot{H}_\mu$ (equivalence classes).

and for the inner products we have

$$(s_1, s_2)_{-1,\mu} := \int_{\mathbb{R}^d} \mathbf{v}_1 \cdot \mathbf{v}_2 \, d\mu = \int_{\mathbb{R}^d} \nabla f_1 \cdot \nabla f_2 \, d\mu,$$

Either of these spaces can be considered as a tangent cone for a measure μ on the Wasserstein space. In [3], $L^2_{\mu,\nabla}(\mathbb{R}^d)$ is used while in [21] the space $s \in \dot{H}^{-1}_{\mu}$ is preferred. Depending on the circumstances, one can be more useful than the other. In the sequel we are going to make use of all of them.

The following theorem describes the relationship between the tangent cone on a point $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and the optimal maps between μ and other points in the Wasserstein space.

Theorem 5.1.3. *(Theorem 8.5.1 in [4]) Let $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, then for the tangent cone $Tan_{\mu}\mathcal{P}_2(\mathbb{R}^d)$ we have*

$$Tan_{\mu}\mathcal{P}_2(\mathbb{R}^d) = \overline{\{\lambda(T^{\star} - I) : \ (I \times T^{\star})_{\#}\mu \in \Pi^{\star}(\mu, T^{\star}_{\#}\mu), \ \lambda > 0\}}^{L^2(\mu;\mathbb{R}^d)}.$$

Remark 5.1.4. *It is important to note that although these tangent cones are defined for every point in the Wasserstein space, they don't actually turn it into a Riemannian manifold. The reason is that not all geodesics can be given with the help of an optimal transportation map. For example, it exists a big class of measures that are not "visible" through the "exponential map", when "sitting" on non-regular measures. For a better understanding of the geometry of the Wasserstein space, one could read [2] or [18].*

Although the tangent cones defined above, do not hold all the information needed to describe the geometry of the Wasserstein space, still can be sufficient for many important applications, since for all absolutely continuous curves a tangent vector exists a.e. (theorem (5.1.1)).

This section is concluded with the following, very nice and useful result by Benamou and Brenier [5] which at this point is a simple corollary of theorem 5.1.1 and the fact that Wasserstein is a constant speed geodesic space:

$$W_2^2(\mu_0, \mu_1) = \min \left\{ \int_0^1 \|\partial_t \mu(t)\|^2_{-1,\mu(t)} \, dt : \mu(0) = \mu_0 \text{ and } \mu(1) = \mu_1 \right\}. \quad (5.1.6)$$

5.2 Subdifferential calculus

Since it is possible to define tangent vectors for every point, it is also possible to define concepts that depend on their existence, like the one of the subdifferential. In the sequel we will restrict to functionals such that have domains satisfying $D(|\partial\mathcal{G}|) \subset \mathcal{P}_{2,r}(\mathbb{R}^d)$.

Definition 5.2.1. *Let $\mathcal{G} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ be a functional with $D(|\partial\mathcal{G}|) \subset \mathcal{P}_{2,r}(\mathbb{R}^d)$ and let also $\mu \in D(|\partial\mathcal{G}|)$. We say that $\mathbf{v} \in L^2(\mu; \mathbb{R}^d)$ belongs to the Fréchet subdifferential $\partial\mathcal{G}$ if*

$$\liminf_{\nu \to \mu} \frac{\mathcal{G}(\nu) - \mathcal{G}(\mu) - \int_{\mathbb{R}^d} < \mathbf{v}(\mathbf{r}), (T_\mu^\nu)^\star(\mathbf{r}) - \mathbf{r} > d\mu}{W_2(\mu, \nu)} \geq 0.$$

In the above definition T_μ^ν is the optimal transportation map from μ to ν.

Definition 5.2.2. *Let $\mathcal{G} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ be a functional with $D(|\partial\mathcal{G}|) \subset \mathcal{P}_{2,r}(\mathbb{R}^d)$. We call $\mathcal{G}$ regular if whenever the subdifferentials $\mathbf{v}_n \in \partial\mathcal{G}(\mu_n)$ satisfy*

$$\begin{cases} \mu_n \to \mu, \ \mathcal{G}(\mu_n) \to l, \ \sup\limits_{n \in \mathbb{N}} \|\mathbf{v}_n\|_{L^2(\mu_n;\mathbb{R}^d)} < \infty \\ \mathbf{v}_n \to \mathbf{v} \ \textit{weakly}, \end{cases} \tag{5.2.1}$$

then $\mathbf{v} \in \partial\mathcal{G}$ and $\mathcal{G}(\mu) = l$.

Theorem 5.2.3. *Let $\mathcal{G} : \mathcal{P}_2(\mathbb{R}^d) \to (-\infty, \infty]$ be a regular functional for which $D(|\partial\mathcal{G}|) \subset \mathcal{P}_{2,r}(\mathbb{R}^d)$ holds. Let also assume that Φ as defined in 2.4.9 admits a minimizer for every τ smaller than some τ'. Then we have that $\mu \in D(|\partial\mathcal{G}|)$ if and only if $\partial\mathcal{G}(\mu)$ ia not empty and*

$$|\partial\mathcal{G}|(\mu) = \inf \left\{ \|\mathbf{v}\|_{L^2(\mu;\mathbb{R}^d)} : \mathbf{v} \in \partial\mathcal{G}(\mu) \right\}. \tag{5.2.2}$$

By the convexity of $\partial\mathcal{G}(\mu)$ there exists a unique vector $\mathbf{v} \in \partial\mathcal{G}(\mu)$ which atains the minimum and it will be denoted with $\partial^\star\mathcal{G}(\mu)$

As it is proven in [4, Lemma 10.1.3] λ convex functionals are always regular. Also for λ convex functionals the following very important theorem holds.

Theorem 5.2.4. *([4, Page 233]) Let $\mathcal{G}$ be a λ convex functional and for an absolutely continuous curve $\mu(t)$, we have*

$$\int_a^b |\partial \mathcal{G}|(\mu(t))|\dot{\mu}|(t)dt < \infty. \tag{5.2.3}$$

Then the map $\mathcal{G}(\mu(t))$ is absolutely continuous and even more the for a.e. $t \in I$ we have

$$\frac{d}{dt}\mathcal{G}(\mu(t)) =< \mathbf{v}_1(t), \mathbf{v}_2(t) >, \tag{5.2.4}$$

where $\mathbf{v}_1(t)$ is the tangent vector of μ at point t and $\mathbf{v}_2 \in \partial\mathcal{G}(\mu(t))$.

5.2.1 Relevant functionals

For $\mu \in \mathcal{P}_2(\mathbb{R}^d)$, we define the *Fisher information*

$$I(\mu) := \begin{cases} \int_{\mathbb{R}^d} \frac{|\nabla\rho(x)|^2}{\rho(x)}\, dx & \text{if } \mu = \rho\, dx \text{ and } \sqrt{\rho} \in H^1(\mathbb{R}^d), \\ \infty & \text{otherwise,} \end{cases} \tag{5.2.5}$$

where $\nabla\rho$ is the distributional derivative of ρ.

Fisher information plays a crucial role in the sequel since it is proved to be equal with $|\partial\mathcal{S}|$ (i.e the norm of the gradient of entropy). More specifically by a slight reformulation of [4, theorem 10.4.13] we have that $|\partial\mathcal{S}|$ is finite if and only if $\sqrt{\rho} \in H^1(\mathbb{R}^d)$ and even more $\partial^\star\mathcal{S}(\mu) = \frac{\nabla\rho}{\rho}$.

We sometimes write $\Delta\mu$ and $\nabla(\mu\nabla\Psi)$ for the functionals in $\mathcal{D}'$ defined by

$$\langle \Delta\mu, f\rangle := \int_{\mathbb{R}^d} \Delta f\, d\mu \qquad \text{and} \qquad \langle \nabla(\mu\nabla\Psi), f\rangle := -\int \nabla\Psi \cdot \nabla f\, d\mu.$$

By using (5.1.4), it is straightforward to see that $\|\Delta\mu\|^2_{-1,\mu} \le I(\mu)$, where the inequality turns to equality when the right hand is finite (Actually it is possible to prove equality in general, but it requires a lot of functional analytic machinery). Similarly we have that $\|\nabla(\mu\nabla\Psi)\|^2_{-1,\mu} \le \int |\nabla\Psi|^2 d\mu$. Here equality holds whenever $\int |\nabla\Psi|^2 d\mu < \infty$, which is certainly true if Ψ satisfies assumptions (6.2.1) and $\mu \in \mathcal{P}_2(\mathbb{R}^d)$. Moreover, as a consequence of the HWI inequality [28, Cor. 20.13], if $\mu \in \mathcal{P}_2(\mathbb{R}^d)$ and $I(\mu) < \infty$ then $\mathcal{S}(\mu) < \infty$.

Lemma 5.2.5. *Let $\Psi \in C^2(\mathbb{R}^d)$ be $\lambda-$convex for some $\lambda \in \mathbb{R}$ and bounded from below. Assume also that $\mu(\cdot)\colon (0,\tau) \to \mathcal{P}_2(\mathbb{R}^d)$ is a Wasserstein-absolutely continuous curve, that satisfies the conditions $\mathcal{E}(\mu(t)), \mathcal{S}(\mu(t)) < \infty \quad \forall t \in [0,\tau]$ and*

$$\int_0^\tau \left(\int_{\mathbb{R}^d} |\nabla \Psi(x)|^2 \mu(t)(x)\, dx + I(\mu(t)) \right) dt < \infty. \tag{5.2.6}$$

Then $t \to \mathcal{F}(\mu(t))$ is absolutely continuous and for a.e $t \in [0,\tau]$ we have

$$\frac{d}{dt}\mathcal{F}(\mu(t)) = \left(\Delta\mu(t) + \nabla(\mu(t)\nabla\Psi), \partial_t \mu(t)\right)_{-1,\mu(t)}. \tag{5.2.7}$$

Proof. This lemma is a direct consequence of [3, Th. 10.3.18]. Since the functional $\mathcal{F}(\mu)$ is lower semicontinuous and λ-geodesically convex, we only need to check condition [3, 10.1.17]. This condition in turn is satisfied by the Cauchy-Schwarz inequality $\langle f, g\rangle_{L^2(a,b)} \leq \|f\|_{L^2(a,b)} \|g\|_{L^2(a,b)}$ and the assumptions. □

CHAPTER 6

LARGE DEVIATIONS OF TRAJECTORIES FOR THE FOKKER-PLANCK EQUATION

.

6.1 Introduction

In this sequel we are going to work with the more general Fokker-Planck equation given by

$$\partial_t \mu(t) = \Delta\mu(t) + div(\mu(t)\nabla\Psi), \quad \mu(0) = \mu_0 \in \mathcal{P}_2(\mathbb{R}^d) \tag{6.1.1}$$

where, $\mathcal{P}_2(\mathbb{R}^d)$ is the set of all measures with finite second moments, Δ is the Laplacian and $\nabla\Psi$ is called the drift (Ψ is the potential). Similar to the diffusion equation, the Fokker-Planck equation has a stochastic counterpart. For every $N \in \mathbb{N}$, let $\mathbf{r}_1, \ldots, \mathbf{r}_N$ be the solutions to the Itō stochastic equations

$$d\mathbf{r}_k(t) = -\nabla\Psi(\mathbf{r}_k(t))\, dt + \sqrt{2}\, d\mathbf{w}_k(t), \quad k = 1, \cdots, N$$

where $\mathbf{w}_1, \ldots, \mathbf{w}_N$ are independent Wiener processes starting from fixed positions $\mathbf{r}_{i,0}$. It is a known result ([14]) that at each $\tau \geq 0$ the *empirical measure* $L_N(\tau)$

[0]The following two chapters are joint work with Michiel Renger and Hong Duong ([15])

converges almost surely in the narrow topology as $N \to \infty$ to the solution of the Fokker-Planck equation (6.1.1) with initial condition μ_0 [14]. Again, similar to the diffusion equation, the rate of convergence is characterized by a LDP. In [20, Prop. 3.2] and [22, Cor. 13], it was proven that the rate functional for the LDP is given by

$$J_\tau(\mu|\mu_0) = \inf \left\{ \mathcal{H}(\xi \| \mu_0 \, P_\tau) : \gamma \in \Pi(\mu_0, \mu) \right\} \tag{6.1.2}$$

where P_τ is now the transition kernel for the Fokker-Planck equation (6.1.1).

In the following chapters, we will not focus just on the empirical measure at time τ but we are going to study the empirical process $L_N(\cdot)$ as a whole. In [11, Th. 4.5] it was proved that if for $\Psi \in C^2(\mathbb{R}^d)$ there is a $C_0 > 0$ such that $|\mathbf{r}||\nabla\Psi(\mathbf{r})| \leq C_0(1+|\mathbf{r}|^2)$ for all $\mathbf{r} \in \mathbb{R}^d$, the empirical process $\{L_N(t)\}_{0\leq t\leq\tau}$ satisfies a large deviation principle in $C([0,\tau], \mathcal{P}(\mathbb{R}^d))$ with good rate functional

$$\tilde{J}_\tau(\mu(\cdot)) = \frac{1}{4}\int_0^\tau \|\partial_t \rho_t - \Delta\mu(t) - \operatorname{div}(\mu(t)\nabla\Psi)\|^2_{-1,\mu(t)} dt, \tag{6.1.3}$$

if the curve $\mu(\cdot)$ is absolutely continuous in the distributional sense; else we set $\tilde{J}_\tau$ to ∞. It follows from a contraction principle [12, Th. 4.2.1] and a change of variables $t \mapsto t/\tau$ that the conditional rate functional (6.1.2) can also be written as

$$J_\tau(\mu|\mu_0) = \inf_{\mu(\cdot)\in C(\mu_0,\mu)} \frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t) - \tau(\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)} \, dt. \tag{6.1.4}$$

In the sequel we will denote with $C(\mu_0,\mu_1), C_{W_2}(\mu_0,\mu_1)$ the set of all weakly continuous, Wasserstein continuous curves respectively, with $\mu : [0,1] \to \mathcal{P}(\mathbb{R}^d)$ and $\mu(0) = \mu_0, \mu(1) = \mu_1$

Remark 6.1.1. *In* (6.1.4) *we implicitly set*

$$\frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t) - \tau(\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)} \, dt = \infty,$$

when the curve is not absolutely continuous in the distributional sense. Therefore, from now on, we shall only consider curves in $C(\mu_0,\mu)$ or $C_{W_2}(\mu_0,\mu)$ that are absolutely continuous in distributional sense.

Observe that the infimum was over narrowly continuous curves. However, we will prove that under the extra assumption that $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ and $\mathcal{F}(\mu_0)$ is finite the infimum can be taken over $C_{W_2}(\mu_0, \mu)$. In this chapter, we characterize a class of potentials Ψ and initial data μ_0 for which (6.1.2) is equal to

$$\inf_{\mu(\cdot)\in C_{W_2}(\mu_0,\mu)} \left\{ \frac{1}{4\tau} \int_0^1 \|\partial_t \mu(t)\|^2_{-1,\mu(t)} \, dt + \frac{\tau}{4} \int_0^1 \|\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi)\|^2_{-1,\mu(t)} \, dt \right.$$
$$\left. + \frac{1}{2}\mathcal{F}(\mu) - \frac{1}{2}\mathcal{F}(\mu_0) \right\}. \tag{6.1.5}$$

6.2 Rewriting the rate functional

Assumption 6.2.1. *Let $\Psi \in C^2(\mathbb{R}^d)$ such that*

1. *Ψ is bounded from below,*
2. *there is a $C_0 > 0$ such that $|\mathbf{r}||\nabla\Psi(\mathbf{r})| \leq C_0(1 + |\mathbf{r}|^2)$ for all $\mathbf{r} \in \mathbb{R}^d$,*
3. *Ψ is $\lambda-$convex for some $\lambda \in \mathbb{R}$,*
4. *there exists constants $0 \leq C_1 < \frac{1}{4}$ and $C_2, C_3 \in \mathbb{R}^+$ such that $|\Delta\Psi(\mathbf{r})| \leq C_1|\nabla\Psi(\mathbf{r})|^2 + C_2\Psi(\mathbf{r}) + C_3$.*

Proposition 6.2.2. *Let $\Psi \in C^2(\mathbb{R}^d)$ satisfy Assumption 6.2.1. Let $\mu_0 \in \mathcal{P}_{2,ac}(\mathbb{R}^d)$ with $\mathcal{F}(\mu_0) < \infty$ and assume $\mu(\cdot) \in C(\mu_0, \mu)$ with $\tilde{J}_\tau(\mu(\cdot))$ finite. Then we have that $\mu(t) \in \mathcal{P}_2(\mathbb{R}^d)$ for every $t \in [0, 1]$ and, furthermore, the curve $\mu(\cdot)$ is absolutely continuous in the Wasserstein sense, and $\mathcal{F}(\mu(t))$ is absolutely continuous with respect to t. Finally, there holds:*

$$\frac{1}{4\tau} \int_0^1 \|\partial_t\mu(t) - \tau(\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)} \, dt = \frac{1}{4\tau} \int_0^1 \|\partial_t\mu(t)\|^2_{-1,\mu(t)} \, dt$$
$$+ \frac{\tau}{4} \int_0^1 \|\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi)\|^2_{-1,\mu(t)} \, dt + \frac{1}{2}\mathcal{F}(\mu(1)) - \frac{1}{2}\mathcal{F}(\mu(0)).$$

Before we prove this theorem we prove two auxiliary lemmas.

Lemma 6.2.3. *Assume that*

1. $\mathcal{E}(\mu_0) < \infty$
2. $\Psi \in C^2(\mathbb{R}^d)$ *satisfies Assumption 6.2.1,*
3. $\mu(\cdot) \in C(\rho_0, \rho)$,
4. $\tilde{J}_\tau(\mu(\cdot)) < \infty$.

Then

$$\int_0^\tau \int_{\mathbb{R}^d} |\nabla\Psi(\mathbf{r})|^2 \mu(t)(d\mathbf{r})\, dt < \infty. \tag{6.2.1}$$

Proof. For simplicity we take $\tau = 1$. We will prove the following statement: there exist $0 < \delta \leq 1$ and $\alpha, \beta > 0$ that depend only on Ψ such that

$$\begin{aligned}\alpha \sup_{t\in[0,\delta]} \int_{\mathbb{R}^d} |\Psi|\, \mu(t)(d\mathbf{r}) + \beta \int_0^\delta \int_{\mathbb{R}^d} |\nabla\Psi|^2\, \mu(t)(d\mathbf{r})\, dt \leq & 8\tilde{J}_1(\mu(\cdot)) + 4\,|\inf\Psi| \\ & + 2\int_{\mathbb{R}^d} \Psi\, d\rho_0 + 2\delta C_3.\end{aligned} \tag{6.2.2}$$

Obviously (6.2.1) follows from (6.2.2) by repeating it $1/\delta$ times.

By [11, Lem. 4.8], for any $0 \leq s \leq 1$ we have

$$\begin{aligned}4\tilde{J}_1(\mu(\cdot)) \geq 4\tilde{J}_s(\mu(\cdot)) = \sup_{f\in C_c^2(\mathbb{R}^d)} & \int_{\mathbb{R}^d} f\, d\rho_s - \int_{\mathbb{R}^d} f\, d\rho_0 - \\ & \int_0^s \int_{\mathbb{R}^d} \left(\Delta f - \nabla\Psi\cdot\nabla f + \frac{1}{2}|\nabla f|^2\right) \mu(t)(d\mathbf{r})\, dt.\end{aligned} \tag{6.2.3}$$

It is worth highlighting that in the above equality, the supremum is taken over $C_c^2(\mathbb{R}^d)$ functions instead of $\mathcal{D}$.

The idea is to use two approximations of Ψ so that it can be chosen as a test function f in (6.2.3). The first approximation is used to show that this inequality still holds if we take replace $C_c^2(\mathbb{R}^d)$ by

$$A := \left\{f \in C^2(\mathbb{R}^d) : f, \nabla f, \Delta f, \mathbf{r}f, |\nabla f|\,|\mathbf{r}| \text{ are all bounded}\right\}. \tag{6.2.4}$$

Take an arbitrary $f \in A$. Define the bump function

$$\zeta(\mathbf{r}) := \begin{cases} \exp\left(1 - \frac{1}{1-|\mathbf{r}|^2}\right), & |\mathbf{r}| < 1, \\ 0, & |\mathbf{r}| \geq 1, \end{cases}$$

and set $\zeta_k(\mathbf{r}) := \zeta(\mathbf{r}/k)$. Then surely $\zeta_k f \in C_c^2(\mathbb{R}^d)$. It is easy to check that

$$|\zeta_k(\mathbf{r})| \leq 1, \quad |\nabla \zeta_k(\mathbf{r})| \leq \frac{1}{k} \quad \text{and} \quad |\Delta \zeta_k(\mathbf{r})| \leq \frac{1}{k^2}. \tag{6.2.5}$$

By the Dominated Convergence Theorem, as $k \to \infty$

$$\begin{aligned}
&\int_{\mathbb{R}^d} \zeta_k f \, d\rho_s \to \int_{\mathbb{R}^d} f \, d\rho_s, \\
&\int_{\mathbb{R}^d} \zeta_k f \, d\rho_0 \to \int_{\mathbb{R}^d} f \, d\rho_0, \\
&\int_0^s \int_{\mathbb{R}^d} \Delta(\zeta_k f) \, \mu(t)(d\mathbf{r}) \, dt = \int_0^s \int_{\mathbb{R}^d} (f \Delta \zeta_k + 2 \nabla \zeta_k \cdot \nabla f + \zeta_k \Delta f) \; \mu(t)(d\mathbf{r}) \, dt \\
&\qquad\qquad\qquad\qquad \to \int_0^s \int_{\mathbb{R}^d} \Delta f \, \mu(t)(d\mathbf{r}) \, dt, \\
&\int_0^s \int_{\mathbb{R}^d} \nabla \Psi \cdot \nabla(\zeta_k f) \, \mu(t)(d\mathbf{r}) \, dt = \int_0^s \int_{\mathbb{R}^d} \nabla \Psi \cdot (f \nabla \zeta_k + \zeta_k \nabla f) \, \mu(t)(d\mathbf{r}) \, dt \\
&\qquad\qquad\qquad\qquad \to \int_0^s \int_{\mathbb{R}^d} \nabla \Psi \cdot \nabla f \, \mu(t)(d\mathbf{r}) \, dt, \\
&\int_0^s \int_{\mathbb{R}^d} |\nabla(\zeta_k f)|^2 \, \mu(t)(d\mathbf{r}) \, dt = \int_0^s \int_{\mathbb{R}^d} |f \nabla \zeta_k + \zeta_k \nabla f|^2 \, \mu(t)(d\mathbf{r}) \, dt \\
&\qquad\qquad\qquad\qquad \to \int_0^s \int_{\mathbb{R}^d} |\nabla f|^2 \, \mu(t)(d\mathbf{r}) \, dt,
\end{aligned}$$

where the absolute finiteness of the right-hand integrals is guaranteed by the properties of the set A. Therefore (6.2.3) indeed becomes

$$4\tilde{J}_1(\mu(\cdot)) \geq \sup_{f \in A} \int_{\mathbb{R}^d} f \, d\rho_s - \int_{\mathbb{R}^d} f \, d\rho_0 - \int_0^s \int_{\mathbb{R}^d} \left(\Delta f - \nabla \Psi \cdot \nabla f + \frac{1}{2} |\nabla f|^2 \right) \mu(t)(d\mathbf{r}) \, dt. \tag{6.2.6}$$

For the second approximation we take

$$\eta(\mathbf{r}) := \exp\left(1 - \sqrt{1 + |\mathbf{r}|^2}\right),$$

and set $\eta_k(\mathbf{r}) := \eta(\mathbf{r}/k)$. Then the following estimates hold

$$|\eta_k(\mathbf{r})| \le 1, \qquad |\nabla\eta_k(\mathbf{r})| \le \frac{1}{k}\eta_k(\mathbf{r}) \qquad \text{and} \qquad |\Delta\eta_k(\mathbf{r})| \le \frac{1}{k^2}\eta_k(\mathbf{r}). \tag{6.2.7}$$

Since $\eta_k\Psi \in A$ by the subquadratic Assumption 6.2.1, we can substitute $\eta_k\Psi$ in (6.2.6):

$$\begin{aligned} 4\tilde{J}_1(\mu(\cdot)) \ge \int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_s - \int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_0 - \int_0^s \int_{\mathbb{R}^d} \Delta(\eta_k\Psi)\, \mu(t)(d\mathbf{r})\, dt \\ + \int_0^s \int_{\mathbb{R}^d} \left(\nabla\Psi\cdot\nabla(\eta_k\Psi) - \frac{1}{2}|\nabla(\eta_k\Psi)|^2\right) \mu(t)(d\mathbf{r})\, dt. \end{aligned} \tag{6.2.8}$$

for any $k \in \mathbb{N}$ and $s \in [0,1]$.

We now estimate each term in the right-hand side of (6.2.8). For the first two terms, we have

$$\int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_s - \int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_0 \ge \int_{\mathbb{R}^d} \eta_k|\Psi|\, d\rho_s - 2\,|\inf\Psi| - \int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_0. \tag{6.2.9}$$

For the third term of (6.2.8), we find

$$\begin{aligned} &-\int_0^s \int_{\mathbb{R}^d} \Delta(\eta_k\Psi)\, \mu(t)(d\mathbf{r})\, dt = -\int_0^s \int_{\mathbb{R}^d} (\Psi\Delta\eta_k + 2\nabla\eta_k\cdot\nabla\Psi + \eta_k\Delta\Psi)\, \mu(t)(d\mathbf{r})\, dt \\ &\ge -\int_0^s \int_{\mathbb{R}^d} \big(|\Delta\eta_k|\,|\Psi| + |\nabla\eta_k|\,(|\nabla\Psi|^2+1) + \eta_k|\Delta\Psi|\big)\, \mu(t)(d\mathbf{r})\, dt \\ &\overset{(6.2.7)}{\ge} -\int_0^s \int_{\mathbb{R}^d} \left(\frac{1}{k^2}\eta_k\Psi + \frac{\eta_k}{k}(|\nabla\Psi|^2+1) + \eta_k|\Delta\Psi|\right) \mu(t)(d\mathbf{r})\, dt \\ &\overset{6.2.1(4)}{\ge} -\int_0^s \int_{\mathbb{R}^d} \left(\frac{1}{k^2}\eta_k\Psi + \frac{\eta_k}{k}(|\nabla\Psi|^2+1)\right) \mu(t)(d\mathbf{r})\, dt \\ &-\int_0^s \int_{\mathbb{R}^d} \big(\eta_k\left(C_1|\nabla\Psi|^2 + C_2|\Psi| + C_3\right)\big)\, \mu(t)(d\mathbf{r})\, dt \\ &\ge -s\left(\frac{1}{k^2} + C_2\right) \sup_{t\in[0,s]} \int_{\mathbb{R}^d} \eta_k|\Psi|\, \mu(t)(d\mathbf{r}) - \\ &\left(\frac{1}{k} + C_1\right) \int_0^s \int_{\mathbb{R}^d} \eta_k|\nabla\Psi|^2\, \mu(t)(d\mathbf{r})\, dt - \frac{s}{k} - sC_3. \end{aligned} \tag{6.2.10}$$

Finally, for the last part of (6.2.8)

$$
\begin{aligned}
&\int_0^s \int_{\mathbb{R}^d} \left(\nabla\Psi \cdot \nabla(\eta_k \Psi) - \frac{1}{2} |\nabla(\eta_k\Psi)|^2 \right) \mu(t)(d\mathbf{r})\, dt \\
&= \int_0^s \int_{\mathbb{R}^d} \left(-\frac{1}{2}|\nabla\eta_k|^2 \Psi^2 + (1-\eta_k)\nabla\eta_k \cdot \Psi\nabla\Psi + (1-\frac{1}{2}\eta_k)\eta_k |\nabla\Psi|^2 \right) \mu(t)(d\mathbf{r})\, dt \\
&\overset{(6.2.7)}{\geq} \int_0^s \int_{\mathbb{R}^d} \left(-\frac{1}{2k^2}\eta_k \Psi^2 - 2\eta_k \Big|\frac{1}{k}\Psi\Big| \Big|\frac{1}{2}\nabla\Psi\Big| + \frac{3}{4}\eta_k|\nabla\Psi|^2 \right) \mu(t)(d\mathbf{r})\, dt, \\
&\geq \int_0^s \int_{\mathbb{R}^d} \left(-\frac{3}{2k^2}\eta_k\Psi^2 + \left(\frac{3}{4} - \frac{1}{4}\right)\eta_k|\nabla\Psi|^2 \right) \mu(t)(d\mathbf{r})\, dt \\
&\geq \int_0^s \int_{\mathbb{R}^d} \left(-\frac{3C_0(1+k^2)}{2k^2}\eta_k|\Psi| + \frac{1}{2}\eta_k|\nabla\Psi|^2 \right) \mu(t)(d\mathbf{r})\, dt \\
&\geq -\frac{3sC_0(1+k^2)}{2k^2} \sup_{t\in[0,s]} \int_{\mathbb{R}^d} \eta_k|\Psi|\, \mu(t)(d\mathbf{r}) + \int_0^s \int_{\mathbb{R}^d} \left(\frac{1}{2}\eta_k|\nabla\Psi|^2 \right) \mu(t)(d\mathbf{r})\, dt,
\end{aligned}
\tag{6.2.11}
$$

where the fourth line follows from Young's inequality, and in the fifth line we used subquadratic Assumption 6.2.1(2). Substituting (6.2.9), (6.2.10) and (6.2.11) into (6.2.8) we get

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \eta_k|\Psi|\, d\mu_s + \int_0^s \int_{\mathbb{R}^d} \frac{1}{2}\eta_k|\nabla\Psi|^2\, \mu(t)(d\mathbf{r})\, dt \\
&\leq 4\tilde{J}_1(\mu(\cdot)) + 2\,|\inf\Psi| + \int_{\mathbb{R}^d} \eta_k\Psi\, d\rho_0 + \frac{s}{k} + sC_3 \\
&+ s\left(\frac{1}{k^2} + C_2 + \frac{3C_0(1+k^2)}{2k^2} \right) \sup_{t\in[0,s]} \int_{\mathbb{R}^d} \eta_k|\Psi|\, \mu(t)(d\mathbf{r}) \\
&+ \left(\frac{1}{k} + C_1 \right) \int_0^s \int_{\mathbb{R}^d} \eta_k|\nabla\Psi|^2\, \mu(t)(d\mathbf{r})\, dt.
\end{aligned}
$$

If we first discard the first term on the left-hand side and maximize the equation over $s \in [0,\delta]$ for some $0 < \delta \leq 1$, then discard the second term and maximise, the sum of the inequalities can be written as

$$\left(1-2\delta\left(\frac{1}{k^2}+C_2+\frac{3C_0(1+k^2)}{2k^2}\right)\right)\sup_{t\in[0,\delta]}\int_{\mathbb{R}^d}\eta_k|\Psi|\,\mu(t)(d\mathbf{r})$$
$$+\left(\frac{1}{2}-\frac{2}{k}-2C_1\right)\int_0^\delta\int_{\mathbb{R}^d}|\nabla\Psi|^2\,\mu(t)(d\mathbf{r})\,dt$$
$$\leq 8\tilde{J}_1(\mu(\cdot))+4\,|\inf\Psi|+2\int_{\mathbb{R}^d}\eta_k\Psi\,d\rho_0+\frac{2\delta}{k}+2\delta C_3.$$

For δ such that $1>2\delta\left(C_2+\frac{3C_0}{2}\right)$, we get $1>2\delta\left(\frac{1}{k^2}+C_2+\frac{3C_0(1+k^2)}{2k^2}\right)$ for suffiently large k, and therefore from Fatou's Lemma

$$\alpha\sup_{t\in[0,\delta]}\int_{\mathbb{R}^d}|\Psi|\,\mu(t)(d\mathbf{r})+\beta\int_0^\delta\int_{\mathbb{R}^d}|\nabla\Psi|^2\,\mu(t)(d\mathbf{r})\,dt$$
$$\leq 8\tilde{J}_1(\mu(\cdot))+4\,|\inf\Psi|+2\int_{\mathbb{R}^d}\Psi\,d\rho_0+2\delta C_3,$$

with $\alpha:=1-2\delta(C_2+\frac{3C_0}{2})>0$ and $\beta:=\frac{1}{2}-2C_1$. The latter is positive by Assumption 6.2.1(4). □

The second auxiliary lemma is:

Lemma 6.2.4. *Let $\epsilon>0$ and $\mu\in\mathcal{P}(\mathbb{R}^d)$ with Lebesgue density ρ be given. Let also $\theta(\mathbf{r}):=\left(\frac{1}{2\pi}\right)^{\frac{d}{2}}e^{\frac{-|\mathbf{r}|^2}{2}}$ be the density of the d-dimensional normal distribution. We define $\theta_\epsilon(\mathbf{r}):=\epsilon^{-d}\theta(\frac{\mathbf{r}}{\epsilon})$ and $\mu_\epsilon:=\mu*\theta_\epsilon$. Then there exists a constant C_ϵ that depends only on ϵ such that $I(\mu_\epsilon)<C_\epsilon$.*

Proof. We have

$$\nabla\rho_\epsilon(\mathbf{r})=(\rho*\nabla\theta_\epsilon)(\mathbf{r})=\int_{\mathbb{R}^d}\rho(\mathbf{r}-\mathbf{y})\nabla\theta_\epsilon(\mathbf{y})\,d\mathbf{y}=-\epsilon^{-2}\int_{\mathbb{R}^d}\rho(\mathbf{r}-\mathbf{y})\mathbf{y}\theta_\epsilon(\mathbf{y})\,d\mathbf{y}.$$

Furthermore

$$|\nabla\rho_\epsilon(\mathbf{r})|^2\leq\epsilon^{-4}\int_{\mathbb{R}^d}\rho(\mathbf{r}-\mathbf{y})|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,d\mathbf{y}\int_{\mathbb{R}^d}\rho(\mathbf{r}-\mathbf{y})\theta_\epsilon(\mathbf{y})\,d\mathbf{y}$$
$$\leq\epsilon^{-4}\rho_\epsilon(\mathbf{r})\int_{\mathbb{R}^d}\rho(\mathbf{r}-\mathbf{y})|\mathbf{y}|^2\theta_\epsilon(\mathbf{y})\,d\mathbf{y}.$$

Now

$$I(\mu_\epsilon) = \int_{\mathbb{R}^d} \frac{|\nabla \rho_\epsilon(\mathbf{r})|^2}{\rho_\epsilon(\mathbf{r})}\, d\mathbf{r} \le \epsilon^{-4} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(\mathbf{r}-\mathbf{y}) |\mathbf{y}|^2 \theta_\epsilon(\mathbf{y})\, d\mathbf{y}\, d\mathbf{r}$$
$$= \epsilon^{-4} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \rho(\mathbf{r}-\mathbf{y})\, d\mathbf{r}\, |\mathbf{y}|^2 \theta_\epsilon(\mathbf{y})\, d\mathbf{y}$$
$$\le \epsilon^{-4} \int_{\mathbb{R}^d} |\mathbf{y}|^2 \theta_\epsilon(\mathbf{y})\, d\mathbf{y} =: C_\epsilon.$$

□

We are now ready to proceed with the

Proof of Proposition 6.2.2. Let $\mu(\cdot)$ satisfy the assumptions of Proposition 6.2.2. By Lemma 6.2.3 we have

$$\int_0^1 \int_{\mathbb{R}^d} |\nabla \Psi(\mathbf{r})|^2 \mu(t)(d\mathbf{r})\, dt < \infty,$$

and therefore

$$\frac{1}{4\tau} \int_0^1 \|\partial_t \mu(t) - \tau \Delta \mu(t)\|^2_{-1,\mu(t)}\, dt$$
$$\le \frac{1}{2\tau} \int_0^1 \|\partial_t \mu(t) - \tau(\Delta \mu(t) + \operatorname{div}(\mu(t) \nabla \Psi))\|^2_{-1,\mu(t)}\, dt$$
$$+ \frac{\tau}{2} \int_0^1 \int_{\mathbb{R}^d} |\nabla \Psi|^2 \mu(t)(d\mathbf{r})\, dt < \infty.$$

Take a $0 < s \le 1$. Since

$$\frac{1}{4\tau} \int_0^s \|\partial_t \mu(t) - \tau \Delta \mu(t)\|^2_{-1,\mu(t)}\, dt < \infty, \tag{6.2.12}$$

we have that $\|\partial_t \mu(t) - \tau \Delta \mu(t)\|^2_{-1,\mu(t)} < \infty$ for almost every t. By [16, Lem. D.34] there is a $\mathbf{v}(t) \in L^2(\mu(t))$ such that

$$\partial_t \mu(t) - \tau \Delta \mu(t) = -\operatorname{div}(\mathbf{v}(t)\, \mu(t))$$

in distributional sense. Take the Gaussian $\theta_\epsilon(\mathbf{r})$ as in Lemma 6.2.4. Then we have

$$\partial_t \mu_{t,\epsilon} - \tau \Delta \mu_{t,\epsilon} = -\operatorname{div}(\mathbf{v}_{t,\epsilon}\, \mu_{t,\epsilon}),$$

where

$$\mu_{t,\epsilon} = \mu(t) * \theta_\epsilon(\mathbf{r}), \qquad \mathbf{v}_{t,\epsilon} = \frac{(\mathbf{v}(t)\,\mu(t)) * \theta_\epsilon(\mathbf{r})}{\mu_{t,\epsilon}}.$$

By [3, Th. 8.1.9] we have

$$\begin{aligned} &\frac{1}{4\tau}\int_0^s \|\partial_t \mu_{t,\epsilon} - \tau\Delta\mu_{t,\epsilon}\|^2_{-1,\mu(t)}\, dt \le \frac{1}{4\tau}\int_0^s \|\mathbf{v}_{t,\epsilon}\|^2_{L^2(\mu_{t,\epsilon})}\, dt \\ &\le \frac{1}{4\tau}\int_0^s \|\mathbf{v}(t)\|^2_{L^2(\mu(t))}\, dt = \frac{1}{4\tau}\int_0^s \|\partial_t\mu(t) - \tau\Delta\mu(t)\|^2_{-1,\mu(t)}\; dt. \end{aligned} \tag{6.2.13}$$

Furthermore by Lemma 6.2.4 we have that

$$\int_0^s \|\Delta\mu_{t,\epsilon}\|^2_{-1,\mu_{t,\epsilon}}\, dt = \int_0^s I(\mu_{t,\epsilon})\, dt \le C_\epsilon, \tag{6.2.14}$$

and therefore

$$\int_0^s \|\partial_t\mu_{t,\epsilon}\|^2_{-1,\mu_{t,\epsilon}}\, dt < \infty. \tag{6.2.15}$$

From (6.2.14) and since $\rho_0 \in \mathcal{P}_2(\mathbb{R}^d)$, by using [16, Lem. D.34] and Lemma 5.1.1 we get that the curve $\mu_{t,\epsilon}$ is absolutely continuous in $\mathcal{P}_2(\mathbb{R}^d)$. In addition, it is straightforward that $\mathcal{S}(\mu_{t,\epsilon})$ is finite for every $0 < t \le s$. From (6.2.14), (6.2.15) and by Lemma 5.2.5, $\mathcal{S}(\mu_{t,\epsilon})$ is absolutely continuous with respect to t. Hence we obtain

$$\begin{aligned} &\frac{1}{4\tau}\int_0^s \|\partial_t\mu_{t,\epsilon} - \tau\Delta\mu_{t,\epsilon}\|^2_{-1,\mu(t)} dt \\ &= \frac{1}{4\tau}\int_0^s \|\partial_t\mu_{t,\epsilon}\|^2_{-1,\mu(t)} dt + \frac{\tau}{4}\int_0^s \|\Delta\mu_{t,\epsilon}\|^2_{-1,\mu(t)} dt - \frac{1}{2}\int_0^s (\Delta\mu_{t,\epsilon}, \partial_t\mu_{t,\epsilon})_{-1,\mu(t)}\; dt \\ &= \frac{1}{4\tau}\int_0^s \|\partial_t\mu_{t,\epsilon}\|^2_{-1,\mu(t)} dt + \frac{\tau}{4}\int_0^s \|\Delta\mu_{t,\epsilon}\|^2_{-1,\mu(t)} + \frac{1}{2}\mathcal{S}(\mu_{s,\epsilon}) - \frac{1}{2}\mathcal{S}(\mu_{0,\epsilon}). \end{aligned}$$

It follows from this and (6.2.13) that

$$\frac{1}{4\tau}\int_0^s \|\partial_t \mu_{t,\epsilon}\|^2_{-1,\mu(t)}dt + \frac{\tau}{4}\int_0^s \|\Delta \mu_{t,\epsilon}\|^2_{-1,\mu(t)}dt + \frac{1}{2}\mathcal{S}(\mu_{s,\epsilon}) - \frac{1}{2}\mathcal{S}(\mu_{0,\epsilon})$$
$$\leq \frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t) - \tau\Delta\mu(t)\|^2_{-1,\mu(t)}dt.$$

Now letting ϵ go to zero and by the lower semicontinuity of the entropy and the Fisher information functionals we get $\mathcal{S}(\mu_s) < \infty$ and $\int_0^s \|\Delta\mu(t)\|^2_{-1,\mu(t)}\,dt < \infty$. Therefore we have

$$\begin{aligned}&\int_0^s \|\partial_t\mu(t)\|^2_{-1,\mu(t)}dt\\ &\leq 2\left(\int_0^s \|\partial_t\mu(t) - \tau\Delta\mu(t)\|^2_{-1,\mu(t)}dt + \tau^2\int_0^s \|\Delta\mu(t)\|^2_{-1,\mu(t)}\,dt\right) < \infty\end{aligned}$$

and

$$\begin{aligned}&\int_0^s \|\Delta\mu(t) + \operatorname{div}\mu(t)\nabla\Psi\|^2_{-1,\mu(t)}dt\\ &\leq 2\left(\int_0^s \|\Delta\mu(t)\|^2_{-1,\mu(t)}\,dt + \int_0^s\int_{\mathbb{R}^d} |\nabla\Psi(\mathbf{r})|^2\mu(t)(\,d\mathbf{r})\,dt\right) < \infty.\end{aligned}$$

By Lemma 5.1.1, the curve $\mu(t)$ is in $AC_{W_2}\left([0,1];\mathcal{P}_2(\mathbb{R}^d)\right)$. Moreover, $t \mapsto \mathcal{F}(\mu(t))$ is absolutely continuous and (5.2.7) holds. Hence we have

$$\begin{aligned}&\frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t) - \tau(\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)}\ dt\\ &= \frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t)\|^2_{-1,\mu(t)}\,dt + \frac{\tau}{4}\int_0^1 \|\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)}\ dt \qquad (6.2.16)\\ &+ \frac{1}{2}\mathcal{F}(\mu_1) - \frac{1}{2}\mathcal{F}(\mu_0).\end{aligned}$$

This finishes the proof of the proposition. □

Now the following is a straightforward result:

Corollary 6.2.5. *Let $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$ with $\mathcal{F}(\mu_0) < \infty$. If $\Psi \in C^2(\mathbb{R}^d)$ satisfies*

Assumption 6.2.1, then

$$J_\tau(\mu|\mu_0) = \inf_{\mu(\cdot)\in C_{W_2}(\mu_0,\mu)} \frac{1}{4\tau}\int_0^1 \|\partial_t\mu(t) - \tau(\Delta\mu(t) + \operatorname{div}(\mu(t)\nabla\Psi))\|^2_{-1,\mu(t)}\,dt.$$

CHAPTER 7

MAIN RESULT, GENERAL CASE

In this chapter the more general result is proven, namely

Theorem 7.0.6. *Let $\mu_0 \in \mathcal{P}_2(\mathbb{R})$ be absolutely continuous with respect to the Lebesgue measure and with density $\rho_0(x)$ being bounded from below by a positive constant in every compact set. Assume that $\int_{\mathbb{R}} |\nabla\Psi(x)|^2\,\mu_0(dx)$ and the Fisher information $I(\mu_0)$ are finite, and that Ψ satisfies Assumption 6.2.1. Then we have*

$$J_\tau(\cdot\,|\mu_0) - \frac{W_2^2(\mu_0,\,\cdot\,)}{4\tau} \xrightarrow[\tau\to 0]{\Gamma} \frac{1}{2}\mathcal{F}(\,\cdot\,) - \frac{1}{2}\mathcal{F}(\mu_0), \quad \text{in} \quad \mathcal{P}_2(\mathbb{R}). \tag{7.0.1}$$

In Theorem 7.1.1 we prove that the Gamma-convergence lower bound holds for any sequence in $\mathcal{P}_2(\mathbb{R})$, equipped with the narrow topology, and in Theorem 7.2.1 we prove the existence of the recovery sequence in the Wasserstein topology. This is equivalent to having Gamma convergence in both topologies.

All theorems in this chapter are valid in higher dimensions except for the existence of the recovery sequence. There are a number of reasons why, at least by the approach of this paper, the argument fails in higher dimensions. First of all, in the proof of Lemma 7.2.3 we use an explicit formula of optimal transport maps in terms of cumulative distribution functions. Secondly, the proof of the same lemma in higher dimensions would require regularity and global estimates of derivatives of the transport map, which are still unknown today (see for example [29, p. 141]).

7.1 Lower bound

In this section we prove the lower bound of the Gamma convergence (7.0.1) in our main result, Theorem 7.0.6.

Theorem 7.1.1 (Lower bound). *Under the assumptions of Theorem 7.0.6, we have for any $\mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$ and all sequences $\mu_1^\tau \in \mathcal{P}_2(\mathbb{R}^d)$ narrowly converging to μ_1*

$$\liminf_{\tau\to 0} \left(J_\tau(\mu_1^\tau|\mu_0) - \frac{W_2^2(\mu_0,\mu_1^\tau)}{4\tau} \right) \geq \frac{1}{2}\mathcal{F}(\mu_1) - \frac{1}{2}\mathcal{F}(\mu_0). \tag{7.1.1}$$

Proof. Take any sequence $\mu_1^\tau \in \mathcal{P}_2(\mathbb{R}^d)$ narrowly converging to a $\mu_1 \in \mathcal{P}_2(\mathbb{R}^d)$. We only need to consider those μ_1^τ for which $J_\tau(\mu_1^\tau|\mu_0) < \infty$. For each such μ_1^τ, by the definition of infimum there exists a curve $\mu^\tau(t) \in C(\mu_0, \mu_1^\tau)$ satisfying

$$\frac{1}{4\tau}\int_0^1 \left\|\partial_t \mu^\tau(t) - \tau(\Delta\mu^\tau(t) + \operatorname{div}(\mu^\tau(t)\,\nabla\Psi))\right\|^2_{-1,\mu^\tau(t)}\,dt \leq J_\tau(\mu_1^\tau|\mu_0) + \tau < \infty. \tag{7.1.2}$$

By Proposition 6.2.2 we have

$$\begin{aligned}
&J_\tau(\mu_1^\tau|\mu_0) + \tau \geq \frac{1}{4\tau}\int_0^1 \left\|\partial_t \mu^\tau(t) - \tau(\Delta\mu^\tau(t) + \operatorname{div}(\mu^\tau(t)\,\nabla\Psi))\right\|^2_{-1,\mu^\tau(t)}\,dt\\
&= \frac{1}{4\tau}\bigg(\int_0^1 \left\|\partial_t\mu^\tau(t)\right\|^2_{-1,\mu^\tau(t)}\,dt + 2\tau(\mathcal{F}(\mu_1^\tau) - \mathcal{F}(\mu_0))\\
&+ \tau^2 \int_0^1 \|\Delta\mu^\tau(t) + \operatorname{div}(\mu^\tau(t)\,\nabla\Psi)\|^2_{-1,\mu^\tau(t)}\,dt\bigg)\\
&= \frac{1}{2}(\mathcal{F}(\mu_1^\tau) - \mathcal{F}(\mu_0)) + \frac{1}{4\tau}\int_0^1 \left\|\partial_t\mu^\tau(t)\right\|^2_{-1,\mu^\tau(t)}\,dt\\
&+ \frac{\tau}{4}\int_0^1 \|\Delta\mu^\tau(t) + \operatorname{div}(\mu^\tau(t)\,\nabla\Psi)\|^2_{-1,\mu^\tau(t)}\,dt\\
&\geq \frac{1}{2}(\mathcal{F}(\mu_1^\tau) - \mathcal{F}(\mu_0)) + \frac{1}{4\tau}\int_0^1 \left\|\partial_t\mu^\tau(t)\right\|^2_{-1,\mu^\tau(t)}\,dt\\
&\geq \frac{1}{2}(\mathcal{F}(\mu_1^\tau) - \mathcal{F}(\mu_0)) + \frac{1}{4\tau}W_2^2(\mu_0,\mu_1^\tau).
\end{aligned}$$

In the last inequality above we have used the Benamou-Brenier formula (5.1.6) for the Wasserstein distance. Finally, using $\mu_1^\tau \to \mu_1$ narrowly with the narrow lower semi-continuity of $\mathcal{F}$, we find that

$$\liminf_{\tau\to 0}\left(J_\tau(\mu_1^\tau|\mu_0)-\frac{W_2^2(\mu_0,\mu_1^\tau)}{4\tau}\right)\geq \frac{1}{2}\mathcal{F}(\mu_1)-\frac{1}{2}\mathcal{F}(\mu_0).$$

□

7.2 Recovery sequence

In this section we prove the upper bound of the Gamma convergence (7.0.1). This will conclude the proof of Theorem 7.0.6.

Theorem 7.2.1 (Recovery sequence)**.** *Under the assumptions of Theorem 7.0.6, for any $\mu_1 \in \mathcal{P}_2(\mathbb{R})$ there exists a sequence $\mu_1^\tau \in \mathcal{P}_2(\mathbb{R})$ converging to μ_1 in the Wasserstein metric such that*

$$\limsup_{\tau\to 0}\left(J_\tau(\mu_1^\tau|\mu_0)-\frac{W_2^2(\mu_0,\mu_1^\tau)}{4\tau}\right)\leq \frac{1}{2}\mathcal{F}(\mu_1)-\frac{1}{2}\mathcal{F}(\mu_0). \tag{7.2.1}$$

The existence of the recovery sequence is proven by making use of the following denseness argument[1]:

Proposition 7.2.2. *Let (X,d) be a metric space and let Q be a dense subset of X. If $\{K_n, n\in\mathbb{N}\}$ and K_∞ are functions from X to $\mathbb{R}$ such that:*

(a) $K_n(q)\to K_\infty(q)$ for all $q\in Q$,

(b) for every $x\in X$ there exists a sequence $q_n\in Q$ with $q_n\to x$ and $K_\infty(q_n)\to K_\infty(x)$,

then for every $x\in X$ there exists a sequence $r_n\in Q$, with $r_n\to x$ such that $K_n(r_n)\to K_\infty(x)$.

Proof. The proof is by a diagonal argument. Take any $x\in X$ and take the corresponding sequence $q_n\to x$ such that $K_\infty(q_n)\to K_\infty(x)$. By assumption, for

[1]A more or less similar idea can be found in [7, Remark 1.29]; Proposition 7.2.2 is slightly stronger.

any $q \in Q$ and $L > 0$ there exists a $n_{L,q}$ such that for any $n \geq n_{L,q}$ there holds $d(K_n(q), K_\infty(q)) < 1/L$. Define

$$l_n := \begin{cases} 1, & 1 \leq n < n_{2,q_2}, \\ 2, & n_{2,q_2} \leq n < \max\{n_{2,q_2}, n_{3,q_3}\}, \\ \dots \end{cases}$$

Take the subsequence $r_n := q_{l_n}$. Observe that $l_n \to \infty$ as $n \to \infty$ such that indeed $q_{l_n} \to x$, and:

$$d(K_n(q_{l_n}), K_\infty(x)) \leq \underbrace{d(K_n(q_{l_n}), K_\infty(q_{l_n}))}_{\leq \frac{1}{l_n}} + d(K_\infty(q_{l_n}), K_\infty(x)) \to 0.$$

□

For a fixed μ_0 satisfying the assumptions of Theorem 7.0.6, we want to apply Proposition 7.2.2 to the situation where

$$\begin{aligned} &X = \mathcal{P}_2(\mathbb{R}), \\ &Q = Q(\mu_0) = \Big\{\mu(x) \in \mathcal{P}_{2,ac}(\mathbb{R}) : \rho(x) \text{ is bounded from below by a positive} \\ &\quad \text{constant in every compact set}, I(\mu), \int_{\mathbb{R}} |\Psi'(x)|^2 \rho(x)\, dx < \infty, \\ &\text{and there exists a} M > 0 \text{such that} \rho_0(x) = \rho(x) \text{for all} |x| > M\Big\}, \\ &K_n(\mu) = J_{\tau_n}(\mu\,|\mu_0) - \frac{W_2^2(\mu_0, \mu)}{4\tau_n}, \\ &K_\infty(\mu) = \frac{1}{2}\mathcal{F}(\mu) - \frac{1}{2}\mathcal{F}(\mu_0). \end{aligned}$$

Assumption *(a)* of Proposition 7.2.2, i.e. pointwise convergence for every $\mu_1 \in Q(\mu_0)$, can be proven as follows. Take $\mu_1 \in Q(\mu_0)$ and let $\mu(t)$ be the geodesic that connects μ_0 and μ_1. In the following Lemma 7.2.3, we will prove that $I(\mu(t))$

and $\int_{\mathbb{R}} |\Psi'(x)|^2 \mu(t)(dx)$ are uniformly bounded, so that we have

$$\int_0^1 \|\partial_t \mu(t) - \tau(\partial_{xx}\mu(t) + \partial_x(\mu(t)\Psi'))\|^2_{-1,\mu(t)}\ dt$$
$$\leq 3\int_0^1 \|\partial_t \mu(t)\|^2_{-1,\mu(t)}\ dt + 3\tau^2 \int_0^1 \|\partial_{xx}\mu(t)\|^2_{-1,\mu(t)}\ dt + 3\tau^2 \int_0^1 \|\partial_x(\mu(t)\Psi')\|^2_{-1,\mu(t)}\ dt$$
$$< \infty.$$

By Proposition 6.2.2 together with Young's inequality:

$$\lim_{\tau\to 0}\left(J_\tau(\mu_1|\mu_0) - \frac{W_2^2(\mu_0,\mu_1)}{4\tau}\right)$$
$$\leq \lim_{\tau\to 0}\left[\frac{\tau}{2}\int_0^1 \left(\int_{\mathbb{R}}\left(\frac{(\mu_t'(x))^2}{\mu(t)(x)} + |\Psi'(x)|^2\mu(t)(x)\right)dx\right)dt\right.$$
$$\left.+\frac{1}{2}\mathcal{F}(\mu_1) - \frac{1}{2}\mathcal{F}(\mu_0)\right] = \frac{1}{2}\mathcal{F}(\mu_1) - \frac{1}{2}\mathcal{F}(\mu_0).$$

The pointwise convergence then follows from this together with the lower bound (7.1.1).

To prove the uniform bounds:

Lemma 7.2.3. *Let $\Psi \in C^2(\mathbb{R})$ with $\Psi(x) > -A - B|x|^2$ for some positive constants. Let $\mu_0 = \rho_0(x)dx \in \mathcal{P}_2(\mathbb{R})$ be absolutely continuous with respect to the Lesbegue measure, where $\rho_0(x)$ is bounded from below by a positive constant in every compact set. Let $\mu_1 \in Q(\mu_0)$ and $\mu(t)$ be the geodesic that connects μ_0 and μ_1. Assume that $\mathcal{E}(\mu_0)$, $I(\mu_0)$ and $\int_{\mathbb{R}} |\Psi'(x)|^2\mu_0(dx)$ are all finite. Then $\mathcal{F}(\mu(t))$, $I(\mu(t))$ and $\int_{\mathbb{R}} |\Psi'(x)|^2\mu_t(\,dx)$ are uniformly bounded with respect to t.*

Proof. Let $T(x)$ be the optimal map that transports μ_0 to μ_1. The geodesic that connects μ_0 and μ_1 is defined by

$$\mu_t(x) = ((1-t)x + tT(x))_\sharp \mu_0(x).$$

First we prove that $I(\mu_t)$ is uniformly bounded with respect to t. In the real line, the map $T(x)$ can be determined via the cumulative distribution functions as follows [29, Sect. 2.2]). Let $F(x)$ and $G(x)$ be respectively the cumulative distribution functions of μ_0 and μ_1, i.e.

$$F(x) = \int_{-\infty}^{x} \rho_0(x)\, dx; \quad G(x) = \int_{-\infty}^{x} \rho_1(x)\, dx.$$

Then $T = G^{-1} \circ F$. We have

$$F(M) + \int_{M}^{+\infty} \rho_0(x)\, dx = G(M) + \int_{M}^{+\infty} \rho_1(x)\, dx = 1. \tag{7.2.2}$$

From (7.2.2) and by the assumption that $\rho_0(x) = \rho_1(x)$ for all $|x| > M$ we find that $F(M) = G(M)$. Hence for all x such that $|x| > M$ we have

$$F(x) = F(M) + \int_{M}^{x} \rho_0(x)\, dx = G(M) + \int_{M}^{x} \rho_1(x)\, dx = G(x).$$

Consequently, for all x with $|x| > M$ we have $T(x) = (G^{-1} \circ F)(x) = x$. Therefore $T'(x) = 1$ for all $|x| > M$. Also since the densities ρ_0, ρ_1 are absolutely continuous (by assumption) we get that $F(x), G(x)$ are differentiable everywhere with $G'(x) = \rho_1(x) > 0$. We deduce that $T(x)$ has a classical derivative everywhere and moreover since $G(T(x)) = F(x)$, by differentiating we get that $T(x)$ satisfies the Monge - Ampère equation.

$$\rho_0(x) = \rho_1(T(x))T'(x).$$

or equivalently (since $\rho_1(x) > 0$),

$$T'(x) = \frac{\rho_0(x)}{\rho_1(T(x))}. \tag{7.2.3}$$

Due to (7.2.3) we have that $T'(x)$ is absolutely continuous and strictly positive, therefore the derivative of T' exists almost everywhere. Now for the derivative of

T' we have

$$\begin{aligned}\frac{T''(x)}{T'(x)} &= (\log(T'(x)))' \\ &= (\log(\rho_0(x)) - \log(\rho_1(T(x)))' \\ &= \frac{\rho_0'(x)}{\rho_0(x)} - \frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}.\end{aligned}$$

Set $T_t(x) = tx + (1-t)T(x)$. For $0 \le t \le 1$ we have

$$\rho_t(x) = \rho_1(T_t(x))T_t'(x), \tag{7.2.4}$$

Since $\rho_1(T_t(x))$ and $T_t'(x)$ are both absolutely continuous so is $\rho_t(x)$. Hence the derivative appeared in (5.2.5) for $I(\mu_t)$ is the classical derivative. Substituting (7.2.4) into (5.2.5) we get

$$\begin{aligned}\int_{\mathbb{R}} \frac{(\rho_t'(x))^2}{\rho_t(x)}\,dx &= \int_{\mathbb{R}} \frac{[(\rho_1(T_t(x))T_t'(x))']^2}{\rho_1(T_t(x))T_t'(x)}\,dx \\ &= \int_{\mathbb{R}} \frac{[\rho_1'(T_t(x))T_t'(x)^2 + \rho_1(T_t(x))T_t''(x)]^2}{\rho_1(T_t(x))T_t'(x)}\,dx \\ &\le 2\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2(T_t'(x))^4}{\rho_1(T_t(x))T_t'(x)}\,dx + 2\int_{\mathbb{R}} \frac{(\rho_1(T_t(x))T_t''(x))^2}{\rho_1(T_t(x))T_t'(x)}\,dx \\ &= 2\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))^3\,dx + 2\int_{\mathbb{R}} \rho_1(T_t(x))\frac{(T_t''(x))^2}{T_t'(x)}\,dx \quad (7.2.5)\end{aligned}$$

Note that in the inequality above we have used $(a+b)^2 \le 2(a^2+b^2)$. To proceed we will estimate each term in the right-hand side of (7.2.5) using the fact that $|T'(x)|$ is bounded and $I(\mu_0), I(\mu_1) < \infty$. For the first part we have

$$\begin{aligned}\int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))^3\,dx &= \int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))(T_t'(x))^2\,dx \\ &\le C^2 \int_{\mathbb{R}} \frac{(\rho_1'(T_t(x)))^2}{\rho_1(T_t(x))}(T_t'(x))\,dx \\ &= C^2 \int_{\mathbb{R}} \frac{(\rho_1'(x))^2}{\rho_1(x)}\,dx \\ &= C^2 I(\mu_1).\end{aligned} \tag{7.2.6}$$

Let B be the ball of radius M centered at the origin. Since $T''(x) = 0$ for all $|x| > M$ we can restrict our calculation for the second part in the ball B.

$$\begin{aligned}
&\int_{\mathbb{R}} \rho_1(T_t(x)) \frac{(T_t''(x))^2}{T_t'(x)}\, dx = \int_B \rho_1(T_t(x)) \frac{(T_t''(x))^2}{T_t'(x)}\, dx \\
&= \int_B \rho_1(T_t(x)) \frac{((1-t)T''(x))^2}{T_t'(x)}\, dx \\
&= \int_B \rho_1(T_t(x)) T_t'(x) \left(\frac{T'(x)(1-t)}{T_t'(x)}\right)^2 \left(\frac{T''(x)}{T'(x)}\right)^2 dx \\
&= \int_B \rho_1(T_t(x)) T_t'(x) \left(\frac{T'(x)(1-t)}{t+(1-t)T'(x)}\right)^2 \left(\frac{\rho_0'(x)}{\rho_0(x)} - \frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}\right)^2 dx \\
&\leq 2\int_B \rho_1(T_t(x)) T_t'(x) \left(\frac{\rho_0'(x)}{\rho_0(x)}\right)^2 dx + 2\int_B \rho_1(T_t(x)) T_t'(x) \left(\frac{\rho_1'(T(x))T'(x)}{\rho_1(T(x))}\right)^2 dx \\
&= 2\int_B \frac{\rho_1(T_t(x))T_t'(x)}{\rho_0(x)} \left(\frac{(\rho_0'(x))^2}{\rho_0(x)}\right) dx \\
&+ 2\int_B \frac{\rho_1(T_t(x))T_t'(x)T'(x)}{\rho_1(T(x))} \left(\frac{(\rho_1'(T(x)))^2}{\rho_1(T(x))} T'(x)\right) dx \\
&\leq C\left(\int_B \frac{(\rho_0'(x))^2}{\rho_0(x)}\, dx + \int_B \frac{(\rho_1'(T(x)))^2}{\rho_1(T(x))} T'(x)\, dx\right) \leq C(I(\mu_0) + I(\mu_1)). \qquad (7.2.7)
\end{aligned}$$

From (7.2.5), (7.2.6) and (7.2.7) we find that

$$I(\mu_t) = \int_{\mathbb{R}} \frac{(\rho_t'(x))^2}{\rho_t(x)}\, dx \leq C(I(\mu_0) + I(\mu_1)).$$

Next we are going to prove the boundedness of the functional $\int_{\mathbb{R}} |\Psi'(x)|^2 \rho_t(x) dx$. Since $T(x) = x$ for $|x| > M$ we have $\rho(t)(x) = \rho_1(x)$ for $|x| > M$. Hence

$$\begin{aligned}
\int_{\mathbb{R}} |\Psi'(x)|^2 \rho_t(x)\, dx &= \int_B |\Psi'(x)|^2 \rho_t(x)\, dx + \int_{|x|>M} |\Psi'(x)|^2 \rho_t(x)\, dx \\
&= \int_B |\Psi'(x)|^2 \rho_t(x)\, dx + \int_{|x|>M} |\Psi'(x)|^2 \rho_1(x)\, dx \\
&\leq C\int_B \rho_t(x)\, dx + \int_{|x|>M} |\Psi'(x)|^2 \rho_1(x)\, dx \\
&\leq C + \int |\Psi'(x)|^2 \rho_1(x)\, dx < \infty.
\end{aligned}$$

Now we repeat the same argument for $\mathcal{E}(\mu_t)$. Finally by [28, Cor. 20.13] we get that $\mathcal{S}(\mu_0), \mathcal{S}(\mu_1)$ are finite and the result for $\mathcal{S}(\mu_t)$ comes from the fact that $\mathcal{S}$ is geodesically convex. $\square$

Finally we prove that for μ_0 satisfying the assumptions in the main Theorem 7.0.6, the set $Q(\mu_0)$ is dense in $\mathcal{P}_2(\mathbb{R})$, thus satisfying assumption *(b)* of Proposition 7.2.2. The idea behind the lemma is a simple modification of a cut and glue argument (see Figure 7.1). For a given measure $\mu_1 \in \mathcal{P}_2(\mathbb{R})$, we construct a measure that is in some sense nice and close to μ_1 in a compact set, and equal to μ_0 outside of it. To do so, we first find an interval such that the contribution of both measures μ_0, μ_1 to the functionals $\mathcal{S}$ and $\mathcal{E}$ is small outside that interval. We cut out the part of μ_0 that lies outside the interval, mollify it to ensure both positivity and smoothness, and then add a quadratic decay to get finiteness of the Fisher information functional[2]. For μ_0 we just keep the tails and add a quadratic decay. The approximating probability measure is then produced by a linear combination of the above constructed measures.

Lemma 7.2.4. *Assume that $\mu_0 \in \mathcal{P}_{2,ac}(\mathbb{R})$, with Lebesgue density ρ_0 bounded from below by a positive constant in every compact set, and $\mathcal{F}(\mu_0)$, $\int |\Psi'|^2 \, d\mu_0$ and $I(\mu_0)$ are all finite. Let $\Psi \in C^2(\mathbb{R})$ satisfy Assumption 6.2.1. Then for any $\mu_1 \in \mathcal{P}_2(\mathbb{R})$ there exists a sequence μ^τ in $Q(\mu_0)$ such that $\mu^\tau \to \mu_1$ in the Wasserstein topology, and $\mathcal{F}(\mu^\tau) \to \mathcal{F}(\mu_1)$.*

Proof. Take a $\mu_1 \in \mathcal{P}_2(\mathbb{R})$ with $\mathcal{E}(\mu_1) < \infty$ (otherwise the construction is trivial). First observe that, because $\int x^2 \rho_1(x) dx$, $\int x^2 \rho_0(x) dx$ and $\mathcal{S}(\mu_0), \mathcal{S}(\mu_1)$ are all finite, $\int |\rho_1(x) \log \rho_1(x)| dx$, $\int |\rho_0(x) \log \rho_0(x)| dx$ are also finite [19, Eq. (15)]. Secondly, $\int |\Psi(x)| \rho_1(x) dx$ and $\int |\Psi(x)| \rho_0(x) dx$ are also finite since Ψ is bounded from below in Assumptions 6.2.1. Therefore, for any $\tau > 0$ there exist Lebesgue points $M_\tau^- < -1$ and $M_\tau^+ > 1$ of ρ_1 such that (to ease notation we assume that $-M_\tau^- = M_\tau^+ =:$

[2] It is easy to check that a linear decay, which would have been a simpler choice, is not enough to keep the Fisher information functional finite.

M_τ)

$$\rho_0(-M_\tau), \rho_1(-M_\tau) < \min\left\{\frac{\tau}{|\Psi(-M_\tau)|}, \frac{\tau}{M_\tau^2}\right\} \tag{7.2.8a}$$

$$\rho_0(M_\tau), \rho_1(M_\tau) < \min\left\{\frac{\tau}{|\Psi(M_\tau)|}, \frac{\tau}{M_\tau^2}\right\} \tag{7.2.8b}$$

$$\int_{|x|>M_\tau} \big(\rho_0(x) + |\rho_0(x)\log\rho_0(x)| + x^2\rho_0(x) + |\Psi(x)|\rho_0(x)\big)\, dx < \tau, \tag{7.2.8c}$$

$$\int_{|x|>M_\tau} \big(\rho_1(x) + |\rho_1(x)\log\rho_1(x)| + x^2\rho_1(x) + |\Psi(x)|\rho_1(x)\big)\, dx < \tau. \tag{7.2.8d}$$

Define a new density by cutting the tails of ρ_1, and mollifying it by the Gaussian θ_t from Lemma 6.2.4:

$$\sigma^\tau := (\rho_1\, \mathbb{1}_{[-M_\tau, M_\tau]}) * \theta_{t_\tau}, \tag{7.2.9}$$

where t_τ is chosen sufficiently small such that

$$\int_{-M_\tau}^{M_\tau} |\sigma^\tau(x) - \rho_1(x)| dx < \tau, \tag{7.2.10a}$$

$$\int_{-M_\tau}^{M_\tau} |\Psi(x)\sigma^\tau(x) - \Psi(x)\rho_1(x)| dx < \tau \tag{7.2.10b}$$

$$\int_{-M_\tau}^{M_\tau} |x^2\sigma^\tau(x) - x^2\rho_1(x)| dx < \tau, \tag{7.2.10c}$$

$$\left| \int_{-M_\tau}^{M_\tau} \big(\sigma^\tau(x)\log\sigma^\tau(x) - \rho_1(x)\log\rho_1(x)\big) dx \right| < \tau, \tag{7.2.10d}$$

$$\sigma^\tau(-M_\tau) < \min\left\{\tfrac{\tau}{|\Psi(-M_\tau)|}, \tfrac{\tau}{M_\tau^2}\right\} \tag{7.2.10e}$$

$$\sigma^\tau(M_\tau) < \min\left\{\tfrac{\tau}{|\Psi(M_\tau)|}, \tfrac{\tau}{M_\tau^2}\right\}, \tag{7.2.10f}$$

$$\sigma^\tau(x) > 0 \quad \text{whenever } |x| \leq M_\tau. \tag{7.2.10g}$$

Observe that property (7.2.10f) is feasible, because $-M_\tau$ and M_τ are Lebesgue points of ρ_1 and

$$\sigma^\tau(M_\tau) \leq (\rho_1 * \theta_{t_\tau})(M_\tau) \qquad \text{and} \qquad \sigma^\tau(-M_\tau) \leq (\rho_1 * \theta_{t_\tau})(-M_\tau).$$

In order to construct a suitable approximating sequence for ρ_1, small intervals around $-M_\tau$ and M_τ are needed where bounds of the type (7.2.8b) and (7.2.10f) still hold. Indeed, because of (7.2.10f) and the continuity of Ψ, there exists $0 < a_\tau < 1$ such that for all $x \in [-M_\tau - a_\tau, -M_\tau + a_\tau]$:

$$\rho_0(x) < \min\left\{\tau, \frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{7.2.11a}$$

$$\sigma^\tau(x) < \min\left\{\tau, \frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{7.2.11b}$$

and for all $x \in [M_\tau - a_\tau, M_\tau + a_\tau]$:

$$\rho_0(x) < \min\left\{\tau, \frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}, \tag{7.2.11c}$$

$$\sigma^\tau(x) < \min\left\{\tau, \frac{\tau}{|\Psi(x)|}, \frac{\tau}{x^2}\right\}. \tag{7.2.11d}$$

Note that by assumption $M_\tau > 1$, so that the two intervals can not overlap.

Now, using these intervals, replace the tails of σ^τ, which were introduced by the mollification, by quadratically decaying tails (see Figure 7.1)

$$\kappa^\tau(x) = \begin{cases} \sigma^\tau(x), & |x| \le M_\tau, \\ \left(\frac{x-M_\tau-a_\tau}{a_\tau}\right)^2 \sigma^\tau(M_\tau), & M_\tau < x < M_\tau + a_\tau, \\ \left(\frac{x+M_\tau+a_\tau}{a_\tau}\right)^2 \sigma^\tau(-M_\tau), & -M_\tau - a_\tau < x < -M_\tau, \\ 0, & |x| \ge M_\tau + a_\tau. \end{cases}$$

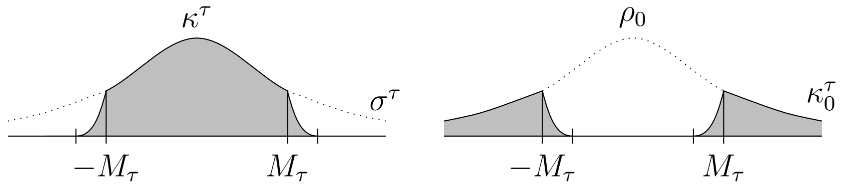
On the other hand, the approximation sequence for μ_1 requires the same tails as μ_0; these tails are captured by (see Figure 7.1)

$$\kappa_0^\tau(x) = \begin{cases} 0, & |x| \le M_\tau - a_\tau, \\ \left(\frac{x-M_\tau+a_\tau}{a_\tau}\right)^2 \rho_0(M_\tau), & M_\tau - a_\tau < x < M_\tau, \\ \left(\frac{x-a_\tau+M_\tau}{a_\tau}\right)^2 \rho_0(-M_\tau), & -M_\tau < x < -M_\tau + a_\tau, \\ \rho_0(x), & |x| \ge M_\tau. \end{cases}$$

Finally, the approximating sequence is defined as a normalized sum of κ and κ_0:

$$\rho^\tau(x) := \alpha_\tau\, \kappa^\tau(x) + \kappa_0^\tau(x), \tag{7.2.12}$$

where $\|\cdot\|_1$ abbreviates the $L^1(\mathbb{R})$ norm, and $\alpha_\tau := \frac{1-\|\kappa_0^\tau\|_1}{\|\kappa^\tau\|_1}$.



(a) Cut the tails, add quadratic decay.

(b) Crop the tails, add quadratic decay.

Figure 7.1: The construction of κ^τ and κ_0^τ.

Now we check that the sequence μ^τ indeed lies in $Q(\mu_0)$. By construction, μ^τ has the same tails as μ_0, and it is bounded from below a positive constant on compact sets. Moreover, it is straight-forward that $\int x^2\, \kappa_0^\tau(x)dx$, $\int x^2\, \kappa^\tau(x)dx$, $\int |\Psi'(x)|^2\, \kappa_0^\tau(x)dx$, $\int |\Psi'(x)|^2\, \kappa^\tau(x)dx$ and, $I(\kappa_0^\tau \mathcal{L})$ are all finite; $I(\kappa^\tau \mathcal{L})$ is finite by Lemma 6.2.4. Then the functionals $\int x^2\, \rho^\tau(x)dx$, $\int |\Psi'(x)|^2\, \rho^\tau(x)dx$ are also finite. To check that the Fisher information remains finite:

$$\begin{aligned}
I(\mu^\tau) &= \int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \kappa^{\tau\prime}(x) + \kappa_0^{\tau\prime}(x)\right)^2}{\alpha_\tau\, \kappa^\tau(x) + \kappa_0^\tau(x)}\, dx \\
&\le 2 \int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \kappa^{\tau\prime}(x)\right)^2}{\alpha_\tau\, \kappa^\tau(x) + \kappa_0^\tau(x)}\, dx + 2 \int_{\mathbb{R}} \frac{\left(\kappa_0^{\tau\prime}(x)\right)^2}{\alpha_\tau\, \kappa^\tau(x) + \kappa_0^\tau(x)}\, dx \\
&\le 2 \int_{\mathbb{R}} \frac{\left(\alpha_\tau\, \kappa^{\tau\prime}(x)\right)^2}{\alpha_\tau\, \kappa^\tau(x)}\, dx + 2 \int_{\mathbb{R}} \frac{\left(\kappa_0^{\tau\prime}(x)\right)^2}{\kappa_0^\tau(x)}\, dx \\
&= 2\alpha_\tau I(\kappa^\tau \mathcal{L}) + 2I(\kappa_0^\tau \mathcal{L}) < \infty,
\end{aligned}$$

so that indeed $\mu^\tau \in Q(\mu_0)$.

Next, the convergence properties of the sequence μ^τ are checked. First we show that $\rho^\tau \to \rho_1$ in $L^1(\mathbb{R})$. Since $\|\kappa_0^\tau\|_1 \to 0$ and $\|\kappa^\tau\|_1 \to 1$, the normalisation constant also converges: $\alpha_\tau \to 1$. Therefore,

$$\begin{aligned}
&\int_{\mathbb{R}} |\rho^\tau(x) - \rho_1(x)| dx = \int_{\mathbb{R}} |\alpha_\tau \kappa^\tau(x) + \kappa_0^\tau(x) - \rho_1(x)|\, dx \\
&\le \int_{-M_\tau}^{M_\tau} |\alpha_\tau \sigma^\tau(x) - \sigma^\tau(x)|\, dx + \int_{-M_\tau}^{M_\tau} |\sigma^\tau(x) - \rho_1(x)| dx + \int_{|x|>M_\tau} \alpha_\tau \kappa^\tau(x) dx \\
&+ \int_{|x|>M_\tau} \rho_1(x) dx + \int_{\mathbb{R}} \kappa_0^\tau(x) dx \le \int_{-M_\tau}^{M_\tau} |\alpha_\tau \sigma^\tau(x) - \sigma^\tau(x)|\, dx \\
&+ \int_{-M_\tau}^{M_\tau} |\sigma^\tau(x) - \rho_1(x)| dx + \alpha_\tau \sigma^\tau(x)(-M_\tau) \int_{-M_\tau - a_\tau}^{-M_\tau} \left(\tfrac{x + M_\tau + a_\tau}{a_\tau}\right)^2 \\
&+ \alpha_\tau \sigma^\tau(M_\tau) \int_{M_\tau}^{M_\tau + a_\tau} \left(\tfrac{x - M_\tau - a_\tau}{a_\tau}\right)^2 + \int_{|x|>M_\tau} \rho_1(x) + \|\kappa_0^\tau(x)\|_1 \\
&\le |\alpha_\tau - 1| \int_{-M_\tau}^{M_\tau} \sigma^\tau(x) dx + \int_{-M_\tau}^{M_\tau} |\sigma^\tau(x) - \rho_1(x)| dx + \alpha_\tau a_\tau\, \sigma^\tau(-M_\tau) \\
&+ \alpha_\tau a_\tau\, \sigma^\tau(M_\tau) + \int_{|x|>M_\tau} \rho_1(x) dx + \|\kappa_0^\tau\|_1 \\
&\le |\alpha_\tau - 1| + \tau + \alpha_\tau \sigma^\tau(-M_\tau) + \alpha_\tau \sigma^\tau(M_\tau) + \tau + \|\kappa_0^\tau\|_1 \xrightarrow[\tau \to 0]{} 0,
\end{aligned} \tag{7.2.13}$$

where the last line follows from $a_\tau < 1$ together with (7.2.10a) and (7.2.8d).

Secondly, we check the convergence of the second moments $\int_{\mathbb{R}} x^2 \rho^\tau(x) dx \to \int_{\mathbb{R}} x^2 \rho_1(x) dx$. Observe that there is a uniform bound on

$$\begin{aligned}
\int_{-M_\tau}^{M_\tau} x^2 \sigma^\tau(x) dx &\le \int_{-M_\tau}^{M_\tau} \left| x^2 \sigma^\tau(x) - x^2 \rho_1(x) \right| dx + \int_{-M_\tau}^{M_\tau} x^2 \rho_1(x) dx \\
&\overset{(7.2.10c)}{<} \tau + \int_{\mathbb{R}} x^2 \rho_1(x) dx \le 1 + \int_{\mathbb{R}} x^2 \rho_1(x) dx
\end{aligned} \tag{7.2.14}$$

for $\tau \le 1$. Moreover, for the right-side quadratic tail of κ^τ:

$$\begin{aligned}
\int_{M_\tau}^{M_\tau + a_\tau} x^2 \kappa^\tau(x) dx &= \int_{M_\tau}^{M_\tau + a_\tau} x^2 \left(\frac{x - M_\tau - a_\tau}{a_\tau}\right)^2 \sigma^\tau(M_\tau)\, dx \\
&\le \int_{M_\tau}^{M_\tau + a_\tau} x^2\, \sigma^\tau(M_\tau)\, dx \overset{(7.2.11d)}{<} \tau\, a_\tau \le \tau,
\end{aligned} \tag{7.2.15}$$

and similarly for the other quadratically decaying parts of κ^τ and κ_0^τ. Therefore

$$\int_{\mathbb{R}} |x^2\rho^\tau(x) - x^2\rho_1(x)|dx \le \int_{\mathbb{R}} |\alpha_\tau x^2 \kappa_\tau(x) - x^2\kappa^\tau(x)|dx$$
$$+ \int_{\mathbb{R}} |x^2\kappa^\tau(x) - x^2\rho_1(x)|dx + \int_{\mathbb{R}} x^2\kappa_0^\tau(x)dx \le |\alpha_\tau - 1| \int_{-M_\tau}^{M_\tau} x^2\sigma^\tau(x)dx$$
$$+ |\alpha_\tau - 1| \int_{|x|>M_\tau} x^2\kappa^\tau(x)dx + \int_{-M_\tau}^{M_\tau} |x^2\sigma^\tau(x) - x^2\rho_1(x)|dx$$
$$+ \int_{|x|>M_\tau} x^2\kappa^\tau(x) + \int_{|x|>M_\tau} x^2\rho_1(x)dx + \int_{-M_\tau}^{M_\tau} x^2\kappa_0^\tau(x)dx + \int_{|x|>M_\tau} x^2\rho_0(x)dx$$
$$\le |\alpha_\tau - 1|\left(1 + \int_{\mathbb{R}} x^2\rho_1(x)dx + 2\tau\right) + 7\tau \to 0$$

as $\tau \to 0$, where the last line follows from (7.2.14), (7.2.15), (7.2.8c), (7.2.8d) and (7.2.10c). Since the sequence ρ^τ converges strongly in $L^1(\mathbb{R})$ to ρ_1 by (7.2.13), it also converges narrowly. Together with the convergence of the second moments, this implies convergence in the Wasserstein distance [29, Th. 7.12], which was to be shown.

Thirdly, we need to check that $\mathcal{E}(\mu^\tau) \to \mathcal{E}(\mu_1)$; this is proven in the same way as the convergence of the second moments above, where x^2 is replaced by the potential $\Psi(x)$.

Finally, we prove the convergence of the entropies $\mathcal{S}(\mu^\tau) \to \mathcal{S}(\mu_1)$. Because of

$$|\mathcal{S}(\mu^\tau) - \mathcal{S}(\mu_1)| \le |\mathcal{S}(\kappa^\tau\mathcal{L}) - \mathcal{S}(\mu_1)| + |\mathcal{S}(\mu^\tau) - \mathcal{S}(\kappa^\tau\mathcal{L})|, \tag{7.2.16}$$

it suffices to show that both differences on the right-hand side vanish. For the first difference:

$$\mathcal{S}(\kappa^\tau\mathcal{L}) = \int_{-M_\tau}^{M_\tau} \big(\kappa^\tau(x)\log\kappa^\tau(x) - \rho_1(x)\log\rho_1(x)\big)dx + \int_{-M_\tau}^{M_\tau} \rho_1(x)\log\rho_1(x)dx$$
$$+ \int_{M_\tau<|x|<M_\tau+a_\tau} \kappa^\tau(x)\log\kappa^\tau(x)dx \to \mathcal{S}(\mu_1).$$

Here, the first term vanishes by (7.2.10d), and the third term, containing the quadratically decaying tails, vanishes because $\sigma^\tau(-M_\tau)$ and $\sigma^\tau(M_\tau)$ vanish. For the second difference in (7.2.16):

$$|\mathcal{S}(\mu^\tau) - \mathcal{S}(\kappa^\tau \mathcal{L})| \leq \underbrace{\int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\alpha_\tau \sigma^\tau(x) \log \alpha_\tau \sigma^\tau(x) - \sigma^\tau(x) \log \sigma^\tau(x)| \, dx}_{(I)}$$
$$+ \underbrace{\int_{M_\tau-a_\tau \leq |x| \leq M_\tau+a_\tau} |\rho^\tau(x) \log \rho^\tau(x) - \kappa^\tau(x) \log \kappa^\tau(x)| \, dx}_{(II)}$$
$$+ \underbrace{\int_{|x|>M_\tau+a_\tau} |\rho_0(x) \log \rho_0(x)| \, dx}_{(III)} .$$

It will now be shown that each of the three parts convergence to 0 as $\tau \to 0$. For the first part, because of (7.2.10d),

$$\begin{aligned}(I) &= \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\alpha_\tau \sigma^\tau(x) \log(\alpha_\tau \sigma^\tau(x)) - \sigma^\tau(x) \log \sigma^\tau(x)| \, dx \\ &\leq |\alpha_\tau - 1| \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\sigma^\tau(x) \log \sigma^\tau(x)| \, dx + \alpha_\tau \log \alpha_\tau \int_{-M_\tau+a_\tau}^{M_\tau-a_\tau} |\sigma^\tau(x)| dx \to 0.\end{aligned}$$

For the second part, observe that by assumptions (7.2.11a), (7.2.11b), (7.2.11c), (7.2.11b), there holds for $M_\tau - a_\tau \leq |x| \leq M_\tau$:

$$\kappa_0^\tau(x) \leq \rho_0(M_\tau) < \tau, \qquad \kappa^\tau(x) = \sigma^\tau(x) < \tau,$$

and for $M_\tau \leq |x| \leq M_\tau + a_\tau$:

$$\kappa_0^\tau(x) = \rho_0(x) < \tau, \qquad \kappa^\tau(x) \leq \sigma^\tau(M_\tau) < \tau.$$

Therefore, since $a_\tau < 1$:

$$(II) \leq \int_{M_\tau - a_\tau \leq |x| \leq M_\tau + a_\tau} (|\rho^\tau(x) \log \rho^\tau(x)| + |\kappa^\tau(x) \log \kappa^\tau(x)|)\, dx =$$
$$\int_{M_\tau - a_\tau \leq |x| \leq M_\tau + a_\tau} |(\alpha_\tau \kappa^\tau(x) + \kappa_0^\tau(x)) \log(\alpha_\tau \kappa^\tau(x) + \kappa_0^\tau(x))|\, dx +$$
$$\int_{M_\tau - a_\tau \leq |x| \leq M_\tau + a_\tau} |\kappa^\tau(x) \log \kappa^\tau(x)| dx \to 0.$$

Finally, part (III) converges to 0 by assumption (7.2.8c). $\square$

7.3 Outlook

As it has been already seen, proving the lower bound for the Gamma convergence, is quite straightforward when someone uses the rate function for the trajectories (7.1.1). However the proof for the upper bound, it may works perfectly for measures on the real line, but it is not at all obvious how it can be generalized in more dimensions. In order to understand the main obstacles, someone has to go through the calculations of lemma 7.2.3. In more dimensions 7.2.3 turns to

$$\det |DT(x)| = \frac{\rho_0(x)}{\rho_1(T(x))}.$$

It is clear that a bootstrap argument cannot be applied since the determinant of the derivative appears instead of the derivative itself. The only known approach will involve looking at the regularity theory for the Monge-Ampére equation. However, even if someone is willing to use such a machinery, will still need an estimate for the growth of T and its derivatives at ∞ (Observe that in 7.2.5 $|DT_t|$ and $\nabla \det(DT_t)$ appears). Although the regularity theory for the Monge-Ampére equation has been expanded significantly during the last decades, first by work of Caffarelli ([8],[9]) and later by Figgali ([17]) and others, it either remains quite local or works for bounded domains, and there is a lack of estimates for the growth of T and of it is derivatives at infinity. In one dimension this problem is remedied by the fact that when two measure agree outside a compact interval, the transport map has to be the identity outside the interval. In more dimensions, the geometry is richer and it is quite difficult to conclude similar properties for the transport map.

BIBLIOGRAPHY

[1] Stefan Adams, Nicolas Dirr, Mark A. Peletier, and Johannes Zimmer. From a Large-Deviations Principle to the Wasserstein Gradient Flow: A New Micro-Macro Passage. *Communications in Mathematical Physics*, 307(3):791–815, September 2011.

[2] L Ambrosio and N Gigli. A user's guide to optimal transport. *Preprint*, 2011.

[3] L Ambrosio, N Gigli, and G Savaré. *Gradient flows: in metric spaces and in the space of probability measures*. Lectures in Mathematics. ETH Zürich. Birkhauser, Basel, 2nd edition, 2005.

[4] Luigi Ambrosio and Wilfrid Gangbo. Hamiltonian ODEs in the Wasserstein space of probability measures. *Communications on Pure and Applied Mathematics*, 000:1–35, 2008.

[5] JD Benamou and Y Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[6] Patrick Billingsley. *Probability and measure*. John Wiley & Sons, 2012.

[7] Andrea Braides. *Gamma-Convergence for Beginners*. Oxford University Press, 2002.

[8] LA Caffarelli. Boundary Regularity of Maps with Convex Potentials–II. *Annals of mathematics*, 144(3):453–496, 1996.

[9] Luis Caffarelli. The Monge-Ampere equation and optimal transportation, an elementary review. *Optimal transportation and applications*, 1(X):1–10, 2003.

[10] Sara Daneri and G Savaré. Lecture notes on gradient flows and optimal transport. *arXiv preprint arXiv:1009.3737*, 2(X):1–27, 2010.

[11] DA Dawson and J Gärtner. Large deviations from the McKean-Vlasov limit for weakly interacting diffusions. *Stochastics*, 20(4):247–308, 1987.

[12] A Dembo and O Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2nd edition, 1987.

[13] Nicolas Dirr, Vaios Laschos, and Johannes Zimmer. Upscaling from particle models to entropic gradient flows. *Journal of Mathematical Physics*, 53(6), 2012.

[14] RM Dudley. *Real analysis and probability*. Cambridge University Press, 2002.

[15] MH Duong, Vaios Laschos, and Michiel Renger. Wasserstein gradient flows from large deviations of thermodynamic limits. *arXiv preprint arXiv:1203.0676*, March 2012.

[16] J Feng and T G Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2006.

[17] A Figalli. Regularity of optimal transport maps (after Ma-Trudinger-Wang and Loeper). *Séminaire Bourbaki*, pages 2008–2009, 2009.

[18] N Gigli. *On the geometry of the space of probability measures endowed with the quadratic optimal transport distance*. PhD thesis, 2004.

[19] Richard Jordan, D Kinderlehrer, and F Otto. The Variational Formulation of the Fokker–Planck Equation. *SIAM Journal on Mathematical Analysis*, pages 1–26, 1998.

[20] C Léonard. A large deviation approach to optimal transport. *arXiv preprint arXiv:0710.1461*, (December):1–35, 2007.

[21] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26(1-2):101–174, January 2001.

[22] M Peletier and M Renger. Variational formulation of the Fokker-Planck equation with decay: a particle approach. *arXiv preprint arXiv:1108.3181*, 2011.

[23] SJ Roberts and WD Penny. Variational Bayes for generalized autoregressive models. *Signal Processing, IEEE Transactions on Signal Processing*, 2002.

[24] Walter Rudin. *Functional Analysis.* McGraw-Hill, 2006.

[25] A Takatsu. Wasserstein geometry of Gaussian measures. *Osaka Journal of Mathematics*, 2011.

[26] John Urbas. Mass transfer problems. *Unpublished*, 1997.

[27] C Villani. Entropy production and convergence to equilibrium. *Entropy Methods for the Boltzmann Equation*, 2008.

[28] C Villani. *Optimal transport: old and new.* Springer, 2008.

[29] Cédric Villani. *Topics in optimal transportation.* American Mathematical Soc., 2003.